Advances in Bi-factor Exploratory Modelling

Eduardo Garcia-Garzon
Universidad Autónoma de Madrid

Advances in Bi-factor Exploratory Modelling

Eduardo Garcia-Garzon

Supervisor: Prof F. J. Abad
Dr L. E. Garrido

Department of Social Psychology and Methodology
Universidad Autónoma de Madrid

This dissertation is submitted for the degree of
Doctor of Philosophy

Universidad Autónoma de Madrid June 2020

"Jet derecha 24. En uno, en uno. Break. Blanco, Blanco 56. Set. Hut.. . . "

Christian García Delgado
Enero, 2017.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university.

Eduardo Garcia-Garzon
June 2020

Acknowledgements

Once I was told that, if you do things right, acknowledgements become the hardest section to be written of your PhD thesis. And today, I can honestly admit that such feeling is nothing but true. In the following pages, I will try to thank everyone who has supported me during these last four years. I must admit this task is never easy when surrounded by such magnificent people. And given my absent-minded character, someone might feel excluded from this text when he or she reads it. Please, take it lightly with me, as these are complicated times. But to those who made it, here are my deepest heartfelt thanks:

Lo primero, agradecer todo su esfuerzo, atención y cariño a mis dos supervisores, Francisco José Abad y Luis Eduardo Garrido. Porque me habéis enseñado a no correr, a (intentar) prestar atención y a ser constante. Aunque llegase un poco sorpresa, siento que estos cuatro años con vosotros han supuesto una etapa inolvidable de mi vida. Me llevo de cada uno mucho más de lo que podéis imaginar. Y ahora que doy mis primeros pasos lejos de vosotros, me doy cuenta de lo mucho que he aprendido en este tiempo. Mi más sincera enhorabuena. No creo que hubiera podido pedir mejores directores de tesis.

Agradecer también al resto de miembros del grupo de investigación y del laboratorio 17 por el tiempo que hemos pasado juntos. Por todos los buenos ratos en nuestras comidas semanales y en nuestros viajes por el mundo. Sé que el laboratorio 17 forma uno de los pilares de la psicometría y la metodología dentro y fuera de nuestro país. Agradeceros la oportunidad de haberme dejado poner mi granito de arena en el mismo. Me llevo el orgullo y el cariño de saber que una parte de mí siempre pertenecerá a la UAM.

Y si hablamos de cariño, cómo olvidarme de toda esa gente del Aula PDIF. A María y Miriam, sois el sol de sur que ha hecho un poco menos oscuro este doctorado. A Jenny, por mudarte de sitio, aguantarme a tu lado todo este tiempo. Por esas listas de reggaetón y por todos dramas. Que no se acaben nunca ni unos, ni otros. A Alba, por aguantar mis idas y venidas. Por hacerlo siempre tener una sonrisa en la cara. Nunca la pierdas. Jamás. Espero que algún día encuentres tu perchero. A María, por luchar conmigo contra mar, cielo y tierra. Porque sin ti, no hubiera sobrevivido estos cuatros años. Por tantos momentos de frustración. Por tantos momentos de alegría. Para Carlos, Lidia, Laura, David, José Ángel, ... por tanto apoyo, cariño y ánimo que me habéis dado. Para mi, sois todos ya catedráticos y catedráticas en lo divino y en lo humano. Jamás podría haber escogido mejores compañeros (que compañeros, ¡amigos!) para pasar estos cuatro años juntos. Aquí os dejo mi último deseo para el árbol de navidad del Aula PDIF: que lo que ha unido el doctorado, nadie lo separe ya.

And whether I would ever make a career in academia, it would all thanks to you, Kai! Your vision not only gave me a chance to discover this world, but encouraged me to continue with this path for the last 7 years. Your commitment and passion taught me not only to be the researcher I am today but further shaped the person I have become. There are not enough words of thanks for all you have done for me. I just hope that someday I will be able to give you back (even in the form of passionate speeches) a little bit of your patience and love!

To everyone involved Junior Researcher Programme. You guys are amazing! Keep on rockin'! To make possible that a simple guy from Salamanca was able to meet such wonderful people. For changing the lifes of countless people. To Matilde, for being such a positive energy blast, for making me smile day in and day out, for celebrating your new job as the head of the United Nations. To the Double Dutchs: To Ondrej, to Cormac, to Dawid and David. For our amazing trips. For the amazing trips to come. For everything we still have to discover together!

And if I talk about the amazing people that JRP let me know, I have to give huge thanks to someone who has been always there. Because you don't know how many times your calls, texts or Yugo-jokes have saved my day. Lea, you are crazy. Crazily amazing. You are one of the most incredible persons I have ever met. I do not care if we live half-a-continent away. I know you are always there. And that's enough for me to know that we will be there for each other for many years to come!

To Neve. For almost two years, you made me felt supported and truly loved. We might have taken different life paths, but I know I would neither be here, nor I would become the person I am today if it wasn't for your support. I could not be happier to have shared a part of my life with you.

Y como no hablar de la otra pasión de mi día a día, Madrid Capitals. Gracias a todos los que habéis formado parte de esta familia, por hacerme amar este deporte. Por hacerme

entender qué es el esfuerzo, qué significa apretar los dientes, qué es luchar y no tener miedo a nada. Por hacerme comprender que un jet derecha 22 es más bonito que un dobles gemelos par derecha verde. Que a esto se juega con fullback o no se juega. Por hacerme entender que aunque la vida te puede llevar a que un ligamento se convierta en dos, y a que llores de rabia por perder un campeonato en el último segundo, también puede traerte experiencias increíbles y gente maravillosa. De esa que haces que vuelvas a entrenar cada día un poquito más fuerte que el anterior. De esa con la que crees que puedes superar todo. Siempre #TimeToMiau.

Y ya que hablamos de arrimar el hombro, es vuestro turno: Mercedes, Pablo, Iñigo. Por estar siempre ahí. Por aguantar los días de malhumor y las noches de trabajo y puerta cerrada. Por cada alegría, por cada tristeza, y por haber fallado miserablemente en mi intento de cocinaros mis famosas lubinas. Por hacer que vivir en Madrid se haya hecho un poquito más agradable. Siempre tendréis una casa, un amigo, un colega, allí donde esté.

A Marta, por hacerme querer Madrid. Por enseñarme a poner las tildes a cómo, dónde, y que este no es esto ni este. Por ver la vida con mirada de niñx y corazón de gigante. Que el mundo sea cada vez un poco más parecido a ti (¡y sin tilde, eh!) y un poco menos parecido al resto. Por no querer cambiarme ni hacerme crecer. Y sin embargo, hacerlo.

Si hablamos de amigos, cómo olvidarme de vosotros, cuadrilla de Psicología. Porque parece mentira que pasan los años, y con vosotros me siento que siempre tengo 20 años. Que podemos estar lejos, pero nunca separados. Esta tesis va también por vosotros. Que aceptásteis a un chaval con perilla que se quería comer el mundo, y le quisisteis. A veces con aciertos aciertos. A veces, con fallos. Por haber hecho que jamás me arrepienta de haber estudiado esta carrera. Sois maravillosos. ¡Y a de kilómetros de viajes para irnos de botellón que nos quedan!

A Marina. Porque seguimos todos esperando a que cambies de opinión y digas que sí.

Para Laura. Esta tesis es especialmente para ti. Tú, que siempre has estado a mi lado. En lo bueno y en lo malo. En lo cercano y en la distancia. Porque sé que el mundo es muy pequeño para ti. Que me muero por saber que será lo siguiente que hagas. Por seguir acompañándote en todos tus viajes. Para que sigamos estando siempre ahí el uno para el otro. Porque me invites de una vez a visitarte a Francia.

A Alberto y Guillermo. A Guillermo y Alberto. Gracias por vuestro apoyo, vuestra cabeza, vuestro consejo. Por todo lo que hemos vivido ya, y todo lo que nos queda por vivir.

Porque ni el municipalismo darwiniano, ni el capitalismo más salvaje, ni la desaparición de la sociedad occidental podrán hacer que dejemos de arreglar el mundo día tras día, mes tras mes, año tras año. Por ser los mejores compañeros que alguien podría pedir. Por las sensaciones de que con vosotros, uno se pasa la vida.

Y si hablamos de amigos, es el momento del GRUPO. A todos y a cada uno de vosotros: Antonio, Cuky, Javi, Jorge, Jose, Loza y Pablo. Daros las gracias. Sois lo jofifo mejor. Sois dips. Sois el secreto que ha hecho posible que hoy esté acabando mi doctorado. ¡Y de que casi no llegue a finalizarlo! Sabiendo que normalmente no dedicamos mucho tiempo a decirnos cosas bonitas, y que a mi me va este rollo, aprovecho estos agradecimientos para escribir, desde el corazón, que se os quiere. Y mucho. Y siempre. Incluso a los engendristas. Sois la suerte de mi vida. Por todo lo que nos queda por recorrer juntos. Por Santa Teresa. Por saber qué vamos a hacer ahora Ciao ponen reggaetón. Esta tesis, y los lapsus mentales que pueda haber, son vuestros también.

A Isabel. Porque escribo estas líneas paralizado. De todas las personas, eres para la única que no tengo palabras. Porque no puedo expresar todo lo que has hecho ya por mi. Y eso dice más que 20 párrafos explicando cada momento en el que me has recogido, me has aupado, y me has animado a continuar escribiendo. No podría reflejar todo lo que has hecho por mí ni en 1000 páginas de agradecimientos. Si hay un final de este camino, es gracias a que tu has estado a mi lado. Y por saber que este final es únicamente el comienzo de algo más grande. De algo mejor. Por ello, darte gracias. Que sigamos celebrando meta a meta, triunfo a triunfo. Pisito de Arganzuela a pisito de Arganzuela. Viaje a Disney a viaje a Disney.

Y ya que las últimas palabras suelen ser las más importantes, quiero reservar el último párrafo de estos agradecimientos para mi familia. Porque muchas veces parece que sois la última prioridad, pero en verdad sois la primera. Sois el factor general que ha estado detrás de mí cada hora, cada día, cada momento de mi vida. No podría haber elegido mejores padre, madre u hermana. Escribo esto con el corazón un poco encogido (porque para algo somos de Castilla, y estas cosas se nos atragantan), pero quiero que sepáis que hoy soy más consciente que nunca de todo lo que habéis hecho por mi. De todo lo que me habéis dado.

Porque hoy me doy cuenta de que durante 30 años, me habéis dado todo.

Esta tesis es lo que puedo devolveros. Es vuestra.

Espero que estéis orgullosos.

Abstract

Bi-factor modelling constitutes today one of the preferred statistical tools in psychological research. Due to its unique characteristics, the bi-factor model has been reintroduced in major areas of interest such as psychopathology, personality or intelligence. Accordingly, our understanding of this model could determine the advancement of our understanding of many psychological phenomena. Despite considerable efforts to appropriately estimate these models, available methods present stringent limitations that question their overall validity and usefulness. This doctoral dissertation aims to provide a detailed account of the historic development of the bi-factor model, as well as its main applications from both, a confirmatory and an exploratory approach. Particularly, this dissertation will detail why and how exploratory bi-factor analysis has emerged as one of the most compelling solutions to approximate these complex structures under realistic settings. Moreover, the central role played by the target rotation will be scrutinized, emphasizing the limitations and potential improvements of current bi-factor target-based rotation methods.

In this context, this doctoral dissertation was ultimately concerned with proposing new alternatives to conduct bi-factor rotation, understanding their benefits with regards to both, parameter estimation and secondary statistics of interest. Furthermore, this thesis dissertation intended to provide free, user-friendly tools aimed for the general public to apply bi-factor exploratory factor analysis. In detail, these objectives were developed in seven chapters:

In Chapter 2, the Iterative Target Rotation based on a Schmid-Leiman solution (i.e., SLi) algorithm was introduced. This algorithm improved the original proposal by Reise, Moore and Maydeu-Olivares (2011) of defining bi-factor target rotation by including Moore, Reise, Depaoli and Haviland (2015) iterative target rotation scheme. Results from a Monte Carlo simulation evidenced that SLi improved factor recovery when compared with widespread methods such as the bi-geomin and the bi-quartimin criteria, the Schmid-Leiman solution or the non-iterative version of the bi-factor target rotation.

In Chapter 3, the Empirical Iterative Target Rotation based on a Schmid-Leiman solution (i.e., SLiD) algorithm was presented. This algorithm aimed to improve the SLi algorithm by including a novel method for the computation of factor-specific, empirically defined cut-off points for distinguishing non-vanishing and vanishing entries in the target matrix. This

strategy was based on finding relevant differences in the distribution of normalized, sorted factor loadings in each group factor. Results from a Monte Carlo simulation demonstrated the superiority of the SLiD algorithm under realistic conditions (i.e., structures presenting a mixture of group factors with different average factor loadings altogether with a high number of cross-loadings).

In Chapter 4, the strategy behind the SLiD algorithm was applied to the study of the properties of the scores from a novel, brief intelligence test: the Last Twelve matrices of the Standard Progressive Matrices (SPM-LS). This application demonstrated the usefulness of the previous strategies to evaluate the assumption of essential unidimensionality as well as the presence of relevant nuisance factors. The application of the methods developed in this thesis revealed that while the scores could be considered essentially unidimensional, the bi-factor model represented the most appropriate measurement model.

In Chapter 5, the consequences of the choice of an exploratory bi-factor rotation were explored with regards to the estimation of the omega hierarchical statistic. Moreover, in this Chapter two new bi-factor rotation methods were studied: the Direct Bi-factor and Direct Schmid-Leiman algorithms. Yet again, results from three different Monte Carlo simulations evidenced that the SLiD algorithm provided the best results regardless structure under consideration (i.e., full, rank-deficient bi-factor models or structures without a general factor), and across a wide range of conditions. Furthermore, it was shown that the application of a partially or completely specified target rotation determined the quality of omega hierarchical estimation when examining target-based algorithms. Lastly, the functioning of each algorithm in eight classical examples was presented to provide a better depiction of the results previously discussed.

In Chapter 6, the integration of the SLiD in the context of exploratory structural equation modelling was discussed. A free, user-friendly Shiny application (*SLiDApp*) was developed so to facilitate the translation of the estimated SLiD target to Mplus. To illustrate the usefulness of this application, a step-by-step guide was introduced using a novel bi-factor examination of the Generic Conspiracionist Belief Scale and its relationship with the Big Five personality traits.

This doctoral dissertation is concluded with a reflective, critical discussion of the benefits and limitations of both, the exploratory bi-factor model and the presented methods for approximating its estimation. In the same spirit, this discussion is focused on presenting future research directions as well as in providing clear advice to applied researchers and psychometricians alike. It is hoped that this dissertation would be helpful to disentangle the benefits and drawbacks of bi-factor modelling, inspiring further research endeavours in this area.

Table of contents

Nomenclature **xvii**

1 Introduction **1**

1.1 Motivation 1

1.1.1 Thesis contributions 2

1.1.2 Thesis Outline 3

1.2 The bi-factor model 4

1.2.1 Mathematical Definition of a Bi-factor Model 5

1.2.2 A Brief Note On Bi-Factor Model Emergence 7

1.2.3 The Impact of Bi-factor Models 8

1.3 Exploratory Factor Models 9

1.3.1 Factor Rotation to Simple Structure 12

1.3.2 Simple Structure in Confirmatory Factor Models 15

1.3.3 Simple Structure in Confirmatory Factor Models 16

1.4 The Case for the Exploratory Bi-factor Model 17

1.4.1 The Schmid-Leiman Orthogonalization 18

1.4.2 A Brief note on Rank Deficiencies and Bi-factor Models 20

1.5 New Methods to approximate Exploratory Bi-factor Model 21

1.5.1 The bi-factor criteria 21

1.5.2 The bi-factor specified target rotation 22

1.6 Challenges on Exploratory Bi-factor Modeling 23

1.6.1 The Fall of the Bi-factor Rotation Criteria 23

1.6.2 Questions in Target Rotation 24

1.7 Bi-factor Models Beyond Factor Loading Recovery 34

1.8 Thesis Contributions and Developments 36

1.8.1 Iterations of Partially Specified Target Matrices: Application to the Bi-factor Case 37

1.8.2 Improving Bi-factor Exploratory Modelling: Bi-factor Rotation based on Loading Differences . . . 37
1.8.3 Searching for G: A New Evaluation of SPM-LS Dimensionality . . 37
1.8.4 On General Factor Reliability: A Comparison of Exploratory Bi-factor Analysis Algorithms . . . 38
1.8.5 Bi-factor Structural Equation Modelling Done Right: An application of the SLiD Algorithm . . . 38

2 Iteration of Partially Specified Target Matrices: Application to the Bi-Factor Case 39

3 Improving Bi-factor Exploratory Modeling 55

4 Searching for G: A New Evaluation of SPM-LS Dimensionality 67

5 On Omega Hierarchical Estimation 87

6 Bi-factor Exploratory Structural Equation Modelling Done Right 107

7 General discussion 125
7.1 Main Results . . . 125
7.1.1 Main Results from Chapter 2 . . . 126
7.1.2 Main Results from Chapter 3 . . . 127
7.1.3 Main Results from Chapter 4 . . . 128
7.1.4 Main Results from Chapter 5 . . . 128
7.1.5 Main Results from Chapter 6 . . . 129
7.2 Future Directions and Limitations . . . 130
7.2.1 The Nature of the Bi-factor Model . . . 130
7.2.2 Questions in Exploratory Bi-factor Model . . . 131
7.2.3 The Expanded Bi-factor Model . . . 132
7.2.4 The Plausibility of the Bi-factor Model . . . 133
7.2.5 Challenges in Factor Rotation . . . 135
7.3 Conclusion . . . 138

References 139

Appendix A Contributed Work 153
A.1 Main Work . . . 153
A.2 Conference: Oral Presentations . . . 153

A.3 Conference: Poster Presentations 154
A.4 Research Awards & Scholarships 154
A.5 Participation in Research Funded Projects 155

Appendix B Resumen 157

Appendix C Discusión General 161
C.1 Resultados Principales 161
C.1.1 Capítulo 2: Principales Resultados 162
C.1.2 Capítulo 3: Resultados Principales 163
C.1.3 Capítulo 4: Resultados Principales 164
C.1.4 Capítulo 5: Resultados Principales 164
C.1.5 Capítulo 6: Resultados Principales 166
C.2 Futuras Direcciones y Limitaciones 166
C.2.1 La Naturaleza del Modelo bi-factorial 166
C.2.2 Preguntas sin Resolver en el Modelo Bi-factorial Exploratorio . . . 168
C.2.3 El Modelo Bi-factorial Expandido 169
C.2.4 La Plausibilidad del Modelo Bi-factorial 170
C.2.5 Nuevos Retos en Rotación Factorial 173
C.3 Conclusiones 176

Nomenclature

Greek Symbols

$\mathbf{Y}$ $I \times J$ matrix of random observed variables

$\mathbf{\Lambda}$ $J \times P$ matrix of fixed factor loadings

$\mathbf{\Phi}$ $P \times P$ matrix of factor covariances

$\mathbf{\Psi}$ $J \times J$ diagonal matrix fixed uniqueness

$\boldsymbol{\zeta}$ $P \times 1$ vector of random factor scores

$\boldsymbol{e}$ $J \times 1$ vector of random errors

Subscripts

$i, \ldots, I$ subject index

$j, \ldots, J$ variable index

$p, \ldots, P$ factor index

Acronyms / Abbreviations

BCFA Bi-factor Confirmatory Factor Analysis

BEFA Bi-factor Exploratory Factor Analysis

CFA Confirmatory Factor Analysis

EFA Exploratory Factor Analysis

ESEM Exploratory Structural Equation Modelling

SEM Structural Equation Modelling

SL Schmid-Leiman Transformation

Chapter 1

Introduction

1.1 Motivation

The bi-factor model is considered today as one of the principal statistical tools for understanding psychological phenomena (Markon, 2019). As of today, the bi-factor model plays a crucial role in major research areas such as intelligence (Molenaar, 2016), psychopathology (Caspi and Moffitt, 2018) or personality (Arias et al., 2018). Despite being proposed more than 80 years ago Holzinger and Swineford (1937), the bi-factor model[1] has remained largely ignored until its recent "rediscovery" (Reise, 2012). Since then, a growing number of bi-factor applications has appeared in the literature, mostly due to its theoretical appeal (Caspi and Moffitt, 2018; Gignac et al., 2017; Markon, 2019), its usefulness for assessing general factors (Zinbarg and Alden, 2015; Zinbarg et al., 2005), and its attractive combination of characteristics from the unidimensional and multidimensional factor models (Chen and Zhang, 2018; Reise, 2012; Reise et al., 2018).

At the same time that the bi-factor model became ubiquitous, psychometricians started to question the appropriateness of the confirmatory version of this factor model. Thus, there has been an increasing interest in exploratory factor solutions during the last decade (Asparouhov and Muthén, 2009; Marsh et al., 2014, 2009). Unsurprisingly, both research trends quickly converged on the first proposals for conducting bi-factor exploratory factor analysis (Jennrich and Bentler, 2011, 2012; Reise et al., 2011). However, these approaches remain today largely unknown (Mansolf and Reise, 2016; Markon, 2019; Reise et al., 2018). In this context, this thesis dissertation was intended to provide a more insightful, thorough exploration of the merits and drawbacks of these bi-factor exploratory factor analysis methods.

[1]Holzinger and Swineford respected the old use of hyphenating the particle "bi" (as in bi-fold). However, as suggested by the Merrian-Webster dictionary, bi could be used add without a hyphen (as in biannual), so many authors currently use the term "bifactor" instead.

1.1.1 Thesis contributions

This thesis dissertation aimed to develop five principal objectives:

1. To delve on the possibilities of the target rotation as a useful tool in the context of BEFA. Building on the initial works of Reise et al. (2011) and Moore et al. (2015), the performance of bi-factor target rotation defined using Schmid-Leiman solution was improved via its iterative application (i.e., the SLi algorithm). This proposal was to be compared with available methods to conduct bi-factor rotation via Monte Carlo simulation.

2. To deepen the understanding of how the target matrix definition could foster or hinder the recovery of a bi-factor exploratory model. Building upon available methods for defining empirical criteria in target rotation (Asparouhov and Muthén, 2009; Fleming, 2003; Jennrich, 2004a; Lorenzo-seva, 1999), a new method for defining bi-factor target matrices models was developed and incorporated within the SLi algorithm.

3. To exemplify the benefits of the proposed methods by analyzing empirical datasets from different areas of interest. Particularly, to explore how BEFA models could help to evaluate the properties of general factors, unravel the presence of group factors, and ascertain the detrimental consequences of an incorrect specification of these structures in general contexts (i.e., SEM). With this objective in mind, the third chapter of this thesis was devoted to illustrating the usefulness of BEFA models in the context of intelligence research.

4. To evaluate how BEFA algorithm choice could determinate the quality of the estimation of relevant secondary statistics such as the omega hierarchical. This statistic was selected for its prominence in the literature, as it has been repeatedly suggested to be a crucial indicator of the strength of a general factor and the quality of a bi-factor solution (Revelle and Condon, 2019; Rodriguez et al., 2015, 2016). Thus, it was of interest to understand what practical consequences could be expected by the selection of a particular BEFA estimation method with regards to the omega hierarchical estimation. Moreover, additional methods for recovering rank-deficient bi-factor models applying a completely specified target rotation were studied for the first time (Giordano and Waller, 2019).

5. To understand the utility of the SLiD algorithm in the context of bi-factor exploratory structural equation modelling (Asparouhov and Muthén, 2009; Marsh et al., 2009). Unfortunately, as current software for performing bi-factor SEM (i.e., *Mplus*) do not

allow researchers to this algorithm, a Shiny to obtain a SLiD-based target matrix was developed. The usefulness of this approach was illustrated by conducting a novel bi-factor SEM exploration of the relationship of the Generic Conspiracionist Believes Scale and the Big Five personality traits.

1.1.2 Thesis Outline

This thesis is organized as follows: *Chapter 1: Introduction* will cover a brief introduction to the bi-factor model, the available mathematical definitions of the confirmatory and exploratory versions of the model and the role of target rotation in recent bi-factor exploratory factor analysis. The second chapter *Chapter 2: Iterations of Partially Specified Target Matrices: Application to the Bi-factor Case* will illustrate the use of the iterative target rotation based on a Schmid-Leiman solution, comparing this approach to previous algorithms available via a Monte Carlo simulation. The third chapter *Chapter 3: Improving Bi-factor Exploratory Modelling: Bi-factor Rotation based on Loading Differences*, will introduce a new algorithm to identify relevant group factor loadings in the context of complex bi-factor models presenting cross-loadings and factors differing in strength. The fourth chapter *Chapter 4: Searching for G: A New Evaluation of SPM-LS Dimensionality* will present an application of BCFA and BEFA models to the validation study of a novel short intelligence test. The fifth chapter *Chapter 5: On Omega Hierarchical estimation: A Comparison of Exploratory Bi-factor Analysis Algorithms* will be focused on assessing the performance of four different BEFA algorithms with regards to their ability to recover general factor reliability when measured using the omega hierarchical. The sixth chapter *Chapter 6: Bi-factor Exploratory Structural Equation Modelling Done Right: Using the SLiDapp Application* will expand the applicability of the SLiD algorithm to ESEM models, presenting a Shiny application (*SLiDApp*) to translate SLiD results from R to *Mplus* software. Lastly, *Chapter 7: Discussion* will introduce a general discussion of the results presented in this thesis alongside future lines of research in the context of exploratory bi-factor modelling.

This dissertation thesis was supervised by Prof Francisco José Abad and Dr Luis Eduardo Garrido, and supported by the Cátedra de Modelos y Aplicaciones Psicométricos de la Universidad Autónoma de Madrid-Instituto de Ingeniería del Conocimiento, the Spanish Ministry of Economy Projects "Test Adaptativos Informatizados Basados en Nuevos Modelos Psicométricos" (PSI2013-44300-P) and "Test Adaptativos Informatizados Multidimensionales: Mejoras en la Calibración y en los Algoritmos de Selección de Items" (PSI2017-85022-P) and financially supported by a PhD grant from the Programa para la Formación del Profesorado Universitario 2015 (FPU 15/03246).

1.2 The bi-factor model

Over a century after being proposed (Spearman, 1904), the common factor model is still regarded as one of the principal statistical tools for understanding psychological phenomena (Mulaik, 2010). The common factor model relies on the *common cause principle*, that argues that latent[2], unobservable *factors* (i.e., intelligence) are the responsible sources of co-variation among observable, measurable variables (i.e., test scores). As such, the common cause principle bridges the assessment of unobservable variables with measures of co-occurrence between variables directly observable in nature. Thus, the common factor model is well-suited to play a substantial role in scientific domains like psychology, whose main objects of interest are, by definition, not directly observable.

The common factor model attributes the co-variation between observable variables to two different sources: (a) latent variables common to sets of the studied variables; (b) unique factors for each observable variable (Mulaik, 2010). To do so, the common factor incorporates three main assumptions (Peeters, 2012; Steiger, 1996):

1. *The Partial Correlation-Explanation rationale*: The covariance between a set of observable variables is explained by the presence of underlying latent factors. In this sense, after conditioning on the latent factors, observable variables should be statistically independent. This property is also known as the *Local Independence* assumption.

2. *The Random Noise rationale*: Common factors are expected to represent systematic sources of variance, while unique factors are set to represent the noise. Similar to other signal-to-noise decomposition models, scores obtained from observable variables will always present an extent of random error.

3. *The True Score rationale*: Following the axioms of the classical test theory (Mcdonald, 1999), common factors and unique factors represent true and error score variance, respectively. The unique factors represent variable-specific and error factors whose effect is not separable under most conventional versions of the factor model.

These principles can be translated into two simple equations: the so-called fundamental equation and the fundamental theorem of factor analysis (Eq.6.12 and 6.13; Mulaik, 2010, p.136). The former defines a set of $i,\dots,I$ realizations from $j,\dots,J$ random variables as a linear function of common and unique factor scores defined in $P \times 1$ ζ and $J \times 1$ e (in the case of a single factor model) such that:

[2] Hereafter, latent variable and common factor have exchangeable meanings.

$$\boldsymbol{Y} = \boldsymbol{\Lambda}\zeta + \boldsymbol{\Psi}e \tag{1.1}$$

where a $J \times P$ $\boldsymbol{\Lambda}$ matrix represent the common factor loading or common factor pattern matrix (i.e., correlations between common factors and variables) and $J \times J$ $\boldsymbol{\Psi}$ represent the unique factor loadings or unique factor pattern matrix. In this classical conceptualization of the common factor model, Y, ζ and e represent random variables, while $\boldsymbol{\Lambda}$ and $\boldsymbol{\Psi}$ are fixed (Stegeman, 2016).

Assuming that $E(\zeta) = 0$, $E[e] = 0$, $E[\zeta\zeta'] = \boldsymbol{I}_{P\times P}$, $E[ee'] = \boldsymbol{I}_{J\times J}$, and $E[\zeta e'] = \boldsymbol{O}_{P\times J}$, being $E[.]$ the mathematical expectation, $\boldsymbol{O}$ the null matrix, $\boldsymbol{I}$ the identity matrix and $'$ the transpose operator, and a little algebra (Mulaik, 2010, p.136), the unique factor scores matrix e can be partialized out. Defining $\boldsymbol{\Phi} = \zeta\zeta'$, the covariance matrix between the observed, random variables ($\boldsymbol{R} = \boldsymbol{Y}\boldsymbol{Y}'$) can be expressed as:

$$\boldsymbol{R} = \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}' + \boldsymbol{\Psi}^2 \tag{1.2}$$

These two equations are the basis of factor analysis. As such, they establish the decomposition of the observed indicators variance-covariance matrix into its common and unique factor parts. To understand why the bi-factor model has gained such prominence in recent years, and how it is different from the traditional unidimensional model, its mathematical properties must be firstly described in detail.

1.2.1 Mathematical Definition of a Bi-factor Model

As of today, there is not a unique mathematical definition of a bi-factor model. While different alternatives have been suggested (Holzinger and Swineford, 1937; Yung et al., 1999), none is yet universally accepted. For the remainder of this doctoral dissertation, the following definitions will apply. A bi-factor model is a multidimensional common P-factor structure (i.e., including assumptions reflected in Eq.1.1 and Eq.1.2) defined by the presence of a single general factor, accounting for common variability between all $j,\ldots,J$ observed variables, plus an additional number of $P-1$ group factors which primarily influence particular subsets of the $j,\ldots,J$ observed variables (Holzinger and Swineford, 1937). These subsets are not restricted to be non-overlapping, but to satisfy, for identification purposes, that at least two (in oblique bi-factor models) or three (in orthogonal bi-factor models) observed variables should present substantial loadings in each group factor (Hayashi and Marcoulides, 2006). Formally, the bi-factor model can be defined, following Eq.1.2 as:

$$\boldsymbol{R} = \boldsymbol{\Lambda}_{BF}\boldsymbol{\Phi}\boldsymbol{\Lambda}_{BF}' + \boldsymbol{\Psi}^2 \tag{1.3}$$

where $\mathbf{\Lambda}_{BF}$ is, as indicated by the subscript BF, in bi-factor form. $\mathbf{\Lambda}_{BF}$ can be though of a block matrix following:

$$\mathbf{\Lambda}_{BF} = \begin{bmatrix} \lambda_G & \mathbf{\Lambda}_{GRP} \end{bmatrix} \tag{1.4}$$

where $\lambda_G = \lambda_{J\times 1}$ represents general factor loadings and $\mathbf{\Lambda}_{Grp} = \mathbf{\Lambda}_{J\times(P-1)}$ represent factor loadings in the group factors. Additionally, the general factor must be orthogonal to other factors in the structure due to identification constraints (Markon, 2019). Therefore, $\mathbf{\Phi}$ could be defined as follows:

$$\mathbf{\Phi} = \begin{bmatrix} 1_{1\times 1} & \boldsymbol{O}_{1\times(P-1)} \\ \boldsymbol{O}_{(P-1)\times 1} & \mathbf{\Phi}_{(P-1)\times(P-1)} \end{bmatrix} \tag{1.5}$$

Noteworthy, the loose structure definition of $\mathbf{\Lambda}$ provided in Eq.1.4 corresponds to an exploratory bi-factor model, where no further restrictions (other than those ensuring its existence, uniqueness and identifiability) are defined. However, the bi-factor model has been traditionally understood from a confirmatory perspective, where additional restrictions are placed: firstly, all $P-1$ group factors are restricted to be orthogonal (i.e., $\mathbf{\Phi} = \boldsymbol{I}_{P\times P}$). Secondly, each $\mathbf{\Lambda}_{GRP_j\cdot}$ row presents, at most, a single, non-zero entry. In other words, each j item loads on a single group factor. Lastly, and in contrast with other common factor models, whether Λ_{BF} should be full-rank remains an open question, with some scholars generalize the name "bi-factor" to structures in which $\mathbf{\Lambda}_G$ could be obtained as a linear combination of $\mathbf{\Lambda}_{GRP}$.

The bi-factor model presents two unique features that distinguish it from alternative multidimensional models: (a) all factors present simultaneous, direct effects on the items; and (b) group factors explain variance not accounted by the general factor (Reise, 2012). This first feature is forced by the particular Λ_{BF} definition, while the second is often a consequence on the extraction method applied plus the specific structure forced in $\mathbf{\Phi}$. The former allows to define the bi-factor as a general model, that under a certain set of restrictions (Yung et al., 1999), could be transformed into a unidimensional, a correlated factor or a higher-order model (Chen and Zhang, 2018; Gignac, 2008; Markon, 2019; Yung et al., 1999)[3]. The later property has fostered the interest in the bi-factor model when investigating the plausibility of a general (and group factors) in dimensionality, fit and test characteristics' assessment.

[3]To bear in mind that the nested relationship between these models could be more complex than expected under many settings (Asparouhov and Muthén, 2019).

1.2.2 A Brief Note On Bi-Factor Model Emergence

Due to the recent interest in bi-factor modelling, its story has been well-documented (Giordano and Waller, 2019), so only a brief historical account of the development of this model would be hereafter presented.

As brilliantly summarized by Mulaik (1986), initial common factor models explored psychological phenomena such as human intelligence using one common factor (Spearman, 1904). For example, a single factor called *g* (later identified with the concept of mental energy by Spearman) was proposed to be responsible for the positive co-variation found in responses to different intelligence-related tasks. However, the presence of a single general factor was quickly questioned, with several authors arguing favouring the study of multiple, correlated common factors (Garnett, 1919; Thomson, 1916). Some authors, including Karl Holzinger and Francis Swineford, went even further and championed the idea of the existence of different layers of common factors to explain the presence of multiple correlated factors. Therefore, the early days of factor analysis saw the emergence of several competing schools of thoughts with regards to the nature of common factors, and thus, different conceptualizations of the psychological attributes (Mulaik, 1986).

The bi-factor model was proposed by Holzinger and Swineford in 1937 as a derivation of the early works regarding Spearmans' unitary conception of intelligence (i.e., the *g* factor; Spearman, 1904). As such, the bi-factor model followed the British school of thought. This was in contrast with the predominant American multiple-factor intelligence theory (Swineford, 1941; Thurstone, 1933)[4]. In this sense, the bi-factor model offered an opportunity for maintaining the privileged position of a single, general mental ability factor while accounting for the presence of multiple, minor factors commonly observed at the time (Holzinger and Swineford, 1937). Nevertheless, as factor analysis research expanded from British to American universities, the interest in solutions including a general factor swiftly diminished in favour of the multiple correlated-factor solutions (Thurstone, 1933, 1940, 1947). Despite their efforts (Swineford, 1941), the predominance of Thurstone's ideas (Mulaik, 1986, 2018) caused the bi-factor model to be disregarded as an anecdotal extension of Spearman's general factor theory (Thurstone, 1947). As a consequence, it was ignored for more than 70 years in the literature (Reise, 2012).

An important landmark in bi-factor modelling was the publication of the Schmid-Leiman transformation (i.e., SL; Schmid & Leiman, 1957). The SL transformation constitutes an approximation to a "bi-factor" model by performing an orthogonal transformation of an exploratory second-order model. Yet again, the SL orthogonalization was largely overshad-

[4]Holzinger and Swineford method for computing bi-factor model is discarded today. However, a fantastic account of their original ideas can be found in Jennrich and Bentler (2011).

owed in the literature. Unfortunately, the SL transformation contributed to the generalization of the incorrect idea that the bi-factor and second-order models could be equivalent (Yung et al., 1999). It was not until the early 90's that the mathematical distinction between both types of structures was again clarified in the literature (Jennsen and Weng, 1994; Mcdonald, 1999; Yung et al., 1999). Particularly, the SL solution was revealed to embed a set of hidden constraints among the vector of general and group factor loadings for items loading in the same group factors that were not expected to occur under the bi-factor model. Accordingly, the higher-order was clarified to represent a restricted version of the bi-factor model (Yung et al., 1999).

After these demonstrations, interest in bi-factor modelling rapidly grew on three main research areas: (a) to understand the usefulness of bi-factor and second-order models for exploring new conceptualization of several psychological constructs (Chen et al., 2006; Reise et al., 2010, 2007); (b) the possibility of substituting the problematic Cronbach's alpha by alternative model-based reliability estimates (such as of omega hierarchical; Zinbarg et al., 2005, 2006); and (c) transcending SL transformation as the principal method for estimating BEFA models as to expand the role of exploratory factor solutions (Jennrich and Bentler, 2011; Reise et al., 2011). These advances in general factor modelling led to the publication of *The Rediscovery of the Bifactor Measurement Models* by Steven P. Reise in 2012. In this seminal article, Reise brilliantly described the strengths and benefits of bi-factor modelling while establishing the foundations for the bi-factor model to become a preferred solution in the factor analysis literature.

1.2.3 The Impact of Bi-factor Models

If Reise intended to stress the usefulness of bi-factor modelling ("the bifactor model [...] provides a strong foundation for understanding psychological constructs and their measurement", Reise, 2012, p.695), he was tremendously successful at it. Since this publication appeared in the literature, interest in bi-factor modelling has exploded (Giordano and Waller, 2019; Markon, 2019; Zhang et al., 2020)[5]. Accordingly, the current impact of bi-factor models in psychological research should not be understated. In a dynamic environment of intense competition between theories and models for understanding psychological phenomena (Borsboom and Wijsen, 2017; Epskamp et al., 2016), the bi-factor model has become ubiquitous in mainstream psychology research areas such as intelligence, personality or psychopathology (Markon, 2019; Reise et al., 2018).

[5]After 2012, the number of publications including a bi-factor model has doubled, at a current rate of near 200 publications per year (Zhang et al., 2020; Figure 1). The citations of the original Schmid-Leiman article citations also showed a similar pattern (Giordano & Waller, 2019, Supplementary Data, Figure 1 and 2).

As an example is better than precept, the ongoing discussions surrounding the general factor of psychopathology (*p-factor*) could help to illustrate the central role that bi-factor models play in certain research areas (Caspi and Moffitt, 2018; Hudziak et al., 2007; Moore et al., 2019; Ronald, 2019; Watt et al., 2019). In this context, the bi-factor model has been extensively used to justify moving onto a hierarchical taxonomy and conceptualization of psychopathology. Whether a general factor of psychopathology is finally accepted as a mainstream theory, the bi-factor model would have played a principal role as the main statistical tool supporting this approach. The potential consequences of accepting a general factor of psychopathology should be underscored: (a) it would change our current understanding regarding how mental disorders appear and evolve; (b) it would request major adjustments on psychological and psychiatric evaluation, treatment and prevention plans are deployed; (c) it would alter how national mental health plans are designed and implemented at the population level; and (d) ultimately, it would dramatically shift the way society and individuals approach mental health. Consequently, moving from our current diagnostic categories to a general factor of psychopathology would indeed impact the lives of future generations of patients and psychologist alike. The consequences can only be thought to be similar to those observed when the general factor *g* was established as the leading theory for explaining human intelligence. Unsurprisingly, these considerations have started to be thoughtfully reflected by major stakeholders in the field (Caspi and Moffitt, 2018; Ronald, 2019).

It is remarkable that while some psychometric applications are expected to have a limited sphere of influence, others could have major personal and societal implications. It should be ensured that these high-stake decisions are based on strong theoretical and statistical foundations. Unfortunately, it might be shocking how little we know today about bi-factor models (Bonifay et al., 2017). Therefore, and to understand the reasons underlying the limited current knowledge regarding these models, a brief historical account regarding the evolution of factor models is presented next.

1.3 Exploratory Factor Models

Factor analysis[6] has traditionally been divided into two distinct classes: *exploratory* and a *confirmatory* factor analysis (Mcdonald, 1999; Mulaik, 2010). In the former, the model is estimated under minimal restrictions. This means that all elements of Eq.1.2 matrices are unrestricted up to identification constraints. The latter approach constitutes a theory-

[6]Factor analysis is a multi-step technique, which requires researchers to conduct a dimensionality determination and to decide on a factor estimation method. The discussion of these steps is beyond the purpose of this thesis dissertation and will be not discussed here.

driven approach, where researchers restrict certain parameters of both matrices to understand whether their prior expectations are supported by the data or not. While the distinction between the "exploratory" and "confirmatory" factor analysis has been accentuated on the theoretical implications of each approach, both approaches only represent two different sets of assumptions within the same, common factor model (Mulaik, 2010). Certain authors go so far to claim that this distinction is ill-advised (Peeters, 2012). Thus, it is important to understand the differences between both approaches to comprehend why confirmatory bi-factor models are being substituted today by their exploratory counterparts (Asparouhov and Muthén, 2009; Marsh et al., 2014, 2009; Reise et al., 2010).

In Eq.1.1 and Eq.1.2 a set of minimal assumptions for deriving the factor model was established. However, the factor model is unidentified in such form[7]. additional restrictions are necessary for the factor model to be identified and to achieve, at least, a certain degree of local uniqueness. As a side note, and for clarity, $\mathbf{\Lambda}$, would be used to refer to a generic, non-bi-factor factor pattern matrix in the next sections. In short, the common factor model presented in 1.2 presents five main identification issues:

1. **Indeterminacy of $\mathbf{\Psi}^2$.** Given Eq.1.2, a condition for factor model existence is that $\boldsymbol{R}-\mathbf{\Psi}^2$ (i.e., the reduced correlation matrix with communalities inserted in its diagonal) is a Gramian matrix with rank equal to $\mathbf{\Lambda}$ rank (Elden and Trendafilov, 2017; Steiger, 2002). Unfortunately, conditions for global identification of the factor model have never been derived (Hayashi and Marcoulides, 2006). This means that given a $\boldsymbol{R}$, multiple different $\mathbf{\Psi}^2$ can be found for $\boldsymbol{R}-\mathbf{\Psi}^2$ to be Gramian and of rank P (i.e., a matrix whose number of nonnegative eigenvalues equal its rank and with the remaining eigenvalues being zero). It is known that given $\boldsymbol{R}$, a factor decomposition is never identified (and therefore, non-unique) if the number of parameters to be estimated is higher to the $J(J+1)/2$ number of non-redundant elements of $\boldsymbol{R}$ (also called the *t-rule*). The model is not identified if $(J-P)^2-J-P<0$ (Anderson and Rubin, 1956; Bekker, 1997; Peeters, 2012). However, this rule is a necessary, not sufficient condition for model identification (Hayashi and Marcoulides, 2006). Nevertheless, while having a large J to P ratio does not ensure that the model is identified, it almost ensures local identifiability under mild conditions. In this sense, many different rules regarding J to P ratios have been deduced, including the famous Lederman bound (Ledermann, 1937) or the Bekker and ten Berge bound (Bekker, 1997), among others (Hayashi and Marcoulides, 2006). Nevertheless, with regards to local identifiability conditions, any

[7]These identification issues are inherent to the common model and most of them remain unresolved if not for recent reformulations of the factor model itself (Adachi and Trendafilov, 2018; Elden and Trendafilov, 2017; Stegeman, 2016).

orthogonal solution presenting a $\mathbf{\Lambda}$ where any column has only two non-zero entries would result in an unidentified (and non-unique) factor model (Anderson and Rubin, 1956). While J to P ratios could approximate factor model identification, they never ensure it. Additionally, and given the factor rotation indeterminacy issue (below), the set of admissible $\mathbf{\Lambda}$ solutions must be restricted to those where each column presents at least three non-zero entries in the orthogonal case and two in the oblique (without further restrictions).

2. **Indeterminacy of $\boldsymbol{\zeta}$**. One the early and stronger criticisms of the factor model was, contrarily to initial Spearman's opinion (Mulaik, 1986), that common and unique factors could not be uniquely identified (Wilson, 1928)[8]. This occurs regardless if $\mathbf{\Lambda}$ or $\mathbf{\Psi}^2$ are identified (see Eq. 13.15 to 13.17 in Mulaik, 2010). This indeterminacy implies that common factors scores could be decomposed into their determinate and indeterminate components, where the latter is not identified given a factor solution. Given a factor decomposition, infinite different $\boldsymbol{\zeta}$ vectors are admissible for that given factor solution. Factor scores in $\boldsymbol{\zeta}$ can only be estimated (Beauducel and Hilger, 2019; Rigdon et al., 2019) after the decomposition has been obtained. The exploration of factor indeterminacy and its consequences, which constitute a topic largely neglected in the literature (Steiger, 1996), has recently gained relevance (Beauducel and Hilger, 2019; Nicewander, 2019; Rigdon et al., 2019). Additionally, resolving this indeterminacy has constituted a motivation to develop refined factor analysis-based decompositions (Adachi and Trendafilov, 2018; Sočan, 2003; Stegeman, 2016).

3. **Idendeterminacy the metric of the common factors**. The common factor metric must be decided by the researcher. When identifying the metric of the factor, a preferred alternative for many years was to fix one $\mathbf{\Lambda}$ entry per column to one. While some authors consider this choice to be inconsequential (Asparouhov and Muthén, 2009), this issue is still unclear (Steiger, 2002). Nevertheless, it is important to bear in mind that these choices are justified by mathematical convenience (Loken, 2005).

4. **Indeterminacy of location $\mathbf{\Phi}$ and $\mathbf{\Lambda}$**. $\mathbf{\Lambda}$ and $\mathbf{\Phi}$ column order is neither defined by the data or model. This problem, akin to label switching issues in mixture models (Peeters, 2012), is of more importance under Bayesian estimation techniques (Fontanella et al., 2019; Peeters, 2012) than in classical factor analysis estimation, where it can be resolved via a matrix projection (Korth and Tucker, 1976).

[8]The 31st volume from *Multivariate behavioural Research* in its 4th Issue includes some classical texts regarding factor analysis indeterminacies from authors as Prof James A. Steiger, Prof. William W. Rozeboom, Prof Roderick P. McDonald or Prof Stanley A. Mulaik, among others, which could not be recommended enough for interested readers.

5. **Rotational Indeterminacy**. Even in the case where $\boldsymbol{\Psi}^2$ would be determinate and unique, $\boldsymbol{\Lambda}$ would suffer of rotational indeterminacy if $P > 1$ (which is always true in the case of $\boldsymbol{\Lambda}_{BF}$). Assuming that a factor decomposition is expected to exist (i.e., $(J-P)^2 - J - P \geq 0$) such as in Eq.1.2, then any non-singular rotation matrix $\mathbf{V}$ can always be found such as:

$$\boldsymbol{R} = \boldsymbol{\Lambda\Phi\Lambda}' + \boldsymbol{\Psi}^2 = (\boldsymbol{\Lambda T})[\boldsymbol{T}^{-1}\boldsymbol{\Phi}(\boldsymbol{T}^{-1})'](\boldsymbol{\Lambda T})' + \boldsymbol{\Psi}^2 = \boldsymbol{\Lambda}^*\boldsymbol{\Phi}^{'*}\boldsymbol{\Lambda}^* + \boldsymbol{\Psi}^2 \tag{1.6}$$

with $\boldsymbol{\Lambda\Lambda'} = \boldsymbol{\Lambda^*\Lambda^{*'}}$ and $\boldsymbol{\Lambda^*} = \boldsymbol{\Lambda T}$ (also true for $\boldsymbol{\Phi}$ and $\boldsymbol{\Phi}*$ matrices), where $\mathbf{T}$ is chosen to each row is unit-norm, so $tr(\boldsymbol{\Phi}) = tr(\boldsymbol{\Phi}^*)$, where *tr* is the trace operator. In other words, $\mathbf{T}$ is selected so $\mathbf{T}$ does not change communality values. This rotational indeterminacy implies that after a factor solution decomposition is found, infinite solutions for $\boldsymbol{\Lambda}$ and $\boldsymbol{\Phi}$ exist, all yielding a similar model fit (Mulaik, 2018).

At this point, two different strategies have been suggested to deal with factor rotation indeterminacy. On one hand, to place a set of minimal restrictions in either $\boldsymbol{\Lambda}$ and $\boldsymbol{\Phi}$, such as the factor model is simultaneously just-identified and become easier to be interpreted. This process is called "factor rotation" (due to the graphical component in the process; Thurstone, 1933; 1947); alternatively, to place as many restrictions on $\boldsymbol{\Lambda}$ and $\boldsymbol{\Phi}$ as possible based on domain knowledge until the factor solution is not only unique but also mathematically overdetermined. This approach was called *confirmatory factor analysis*. In the end, to resolve factor rotation indeterminacy lies indeed in the heart of the distinction between exploratory and confirmatory factor models[9].

1.3.1 Factor Rotation to Simple Structure

It is important to understand how rotational indeterminacy was intrinsically related to the emergence of CFA techniques (Mulaik, 1986). Following Eq.1.2, if $\boldsymbol{\Phi}$ would present diagonal entries of value one corresponding to factor variances, then P constraints are set the model. Alongsie the definition of independent error variances (i.e., $\boldsymbol{\Psi}^2$ being restricted to be a diagonal matrix), a minimal number $P(P-1)/2$ restrictions are left to be set in either $\boldsymbol{\Lambda}$ or non-diagonal $\boldsymbol{\Phi}$ elements. The first proposal to handle rotation indeterminacy was to set a minimal set of restrictions for the model to be identified and to take advance of those restrictions to make the model more interpretable (Thurstone, 1947). This procedure

[9]An approach strongly supported by this thesis author would be to just avoid the distinction between exploratory and confirmatory factor models and rather consider these models within the unrestricted-restricted continuum of factor solutions, as suggested by Ferrando and Lorenzo-Seva (2000) 20 years ago.

was called factor rotation. Factor rotation is performed via a transformation matrix **T**. In the orthogonal case, **T** is chosen so $\mathbf{\Phi}^2 = \mathbf{T'T} = \mathbf{I}$, implicitly imposing the necessary $P(P-1)/2$ constraints. In the oblique case, **T** is chosen so it meets P constraints such as $Diag(\mathbf{\Phi}^2) = Diag(\boldsymbol{T}^{-1}(\boldsymbol{T}^{-1\prime})) = \mathbf{I}$, being *Diag* the diagonal operator that refers to the diagonal elements of a matrix. As infinite **T** exist that will comply with these requirements, the only decision left is to select one of the possible **T** from that set (Browne, 2001).

Thurstone (1933, 1947) proposed the concept of the simple structure as a way of selecting an adequate rotation matrix. The concept of simple structure quickly became one of the central concepts in factor analysis (Browne, 2001; Ertel, 2011; Mulaik, 2010). Thurstone comprehended that a researcher could take advantage of the rotation indeterminacy as to approximate a simplest, more interpretable and reproducible factor solution[10]. Originally suggested in *The Vectors of Mind* (Thurstone, 1933), the concept of simple structure represents a set of rules suggested to help researchers to obtain parsimonious (an thus, interpretable) factor solutions via a factor rotation. Noteworthy, the simple structure was expected to ensuring that the factor solution was of theoretical meaning, and not by only defined by statistical convenience: "[the simple structure concept] was arrived at by psychological considerations and not by any statistical reasoning" (Thurstone, 1940, p.193). By fostering factor parsimony, Thurstone saw the concept of simple structure as a method for revealing common factors with an "objective status not depending on anyone sample of [the] variables of the domain" (Mulaik, 2010, p.279). Accordingly, the ultimate goal of the simple structure was to ensure factor replicability across different sets of tests (Catell, 1966; Kaiser, 1958).

The simple structure concept was initially applied to the task of finding simple reference vectors via a performing graphical factor rotation. Reference vectors were crucial in the early days of factor analysis, as they enabled early researchers to avoid working with the inverse of the rotation matrix in the oblique rotation problem. However, issues regarding the translation between finding a matrix of simple reference vectors and its corresponding simple $\mathbf{\Lambda}$ were notable (Harman, 1967; Mulaik, 2010). Rotation methods using reference structures (called indirect rotations) were quickly abandoned as Jennrich and Sampson (1966) developed a method for the direct rotation of $\mathbf{\Lambda}$ (i.e., the so-called direct rotation methods). Since then, simple structures are defined in terms of factor loadings simplicity. Thurstone's original rules, expressed in terms of $\mathbf{\Lambda}$ simplicity, would be (Browne, 2001):

1. Each row should have at least one zero.

2. For each column, there should be a distinct set of P linearly independent indicators whose factor loading λ_{jp} are zero.

[10] Alternative approaches such as restricting $\boldsymbol{\Lambda\Lambda^T}$ or $\boldsymbol{\Lambda^T\Psi^{-2}\Lambda}$ to be diagonal produce solutions that are seldom interpretable (Jöreskog, 1977).

3. For each pair of columns, there should be several indicators whose entries λ_{jp} vanish in one column but not in the other.

4. For every pair of columns, a large proportion of the indicators should have zero entries in both columns. This applies to factor problems with four or more common factors.

5. For every pair of columns, there should preferably be only a small number of indicators with non-vanishing entries in both columns.

Unfortunately, the translation of these rules to mathematical rules is unclear (McDonald, 1984; Yamashita and Adachi, 2019; Yates, 1987)[11]. Noteworthy, even though most authors have focused on the fifth rule (Bandalos, 2018), only the first condition is concerned with model identifiability, with the remaining set to improve factor replicability (Browne, 2001; Yates, 1987). Either way, since the proposal of the simple structure, countless rotational strategies were developed to approximate these structures, a lot of times without explicit rules for their application or without clear advantages over available procedures (Browne, 2001; Fleming, 2012; Jennrich, 2007). As a side note, factor rotation continues being poorly understood by many researchers and applied based on false premises and software default options (Browne, 2001; Izquierdo et al., 2014; Sellbom and Tellegen, 2019).

Thurstone's works became the cornerstone of the American school of factor analysis, guiding the development of factor analysis research for more than a century[12]. In this sense, the rotation problem established itself as of the principal issues in factor analysis research: "The rotational problem is one of the most important in factor analysis" (Thurstone, 1947, p.108). Nevertheless, this prominence was at the cost of the study of other indeterminacies of factor model (Steiger, 1996). Furthermore, and as Thurstone envisioned, rotation indeterminacy was meant to be used to researcher's advantage as to obtain a simple solution (i.e., the "rotational freedom"; Yamashita & Adachi, 2019). Accordingly, the study of factor rotation was crucially linked to the estimation of interpretable and reproducible factor solutions, which was the ultimate objective of factor analysts at the time. The rotation indeterminacy issue was no longer considered as a pure mathematical, technical issue (i.e., ensuring rotational uniqueness) but to be on the spotlight of factor analysis research during the ensuing decades (Ertel, 2013). In the end, this situation would ultimately result in the development of confirmatory factor analysis as the final solution to obtain simple structures

[11]For a modern translation of the rules into objective, mathematical definitions, see Table 1, Yamashita & Adachi (2019).

[12]Thurstone's Multiple Factor Analysis remains today an indispensable work that should be read by any researcher aspiring to work on this topic. As a tribute, the diagram represented in this doctoral thesis cover was inspired in the bi-factor representation appearing in Figure 6a (p.188).

without the burden of factor rotation (Anderson and Rubin, 1956; Asparouhov and Muthén, 2009; Millsap, 2001; Mulaik, 1986).

1.3.2 Simple Structure in Confirmatory Factor Models

Despite their early success (Carroll, 1953), with many different rotation procedures being proposed in the literature [13], it was quickly acknowledged that mechanic rotations rarely worked to perfection (Ertel, 2013; Yates, 1987). As such, and better expressed by Howe (1955) with regards to Carrolls's Quartimin and other mechanical rotations: "the general consensus seems to be that at the present time, graphical or other methods calling for human judgement, are better" (p.46). Particularly, the idea of $\mathbf{\Lambda}$ and $\mathbf{\Phi}$ values being dependent upon rotation choice made many researchers uncomfortable (Catell, 1966; Hurley and Cattell, 1962). In this spirit, even rotation procedures that would imitate the human judgement would be developed, such as maxplane (Cattell and Muerle, 1960). In the end, the interest shifted from purely exploratory factor analysis to confirmatory alternatives, namely the target rotation and the confirmatory factor analysis (Cattell, 1978; Mulaik, 1986). The latter started with authors argued that "a priori" information should be included in the factor analyst process (as called in Anderson and Rubin, 1956, p.132). In this context, the works of Howe (1955) and Anderson and Rubin (1956) clarified a set of necessary, but again not sufficient conditions for establishing factor solution identification based on imposed additional constraints to those proposed by the exploratory factor model (Dun, 1973). Those included:

1. $\mathbf{\Phi}$ to be a diagonal matrix, which ensures its positive definitiveness.
2. That a bare minimum of $P-1$ zeroes should be fixed in each $\phi\Lambda$ column.
3. The rank of $\mathbf{\Lambda}_p$, defined as the matrix retaining rows whose entries have been fixed to zero in a column p with these zeroes deleted, must be of value $p-1$ for all $p = 1, \ldots, P$ (Theorem 5.7, Anderson & Rubin, 1956).

Thus, fixing $\mathbf{\Lambda}$ values to zero was presented as an optimal strategy that made factor rotation redundant in the factor analysis process. Furthermore, it also made possible to obtain even simpler structures than those obtained following Thurstone's approach. In Jöreskog's words:

"to resolve the problem of rotation, Thurstone proposed the concept of simple structure [...]. The general idea of simple structure is that if the factor has real psychological meaning,

[13] As Mulaik (1986) commented: "resolving the rotation problem became a kind of Holy Grail for factor analyst to pursue, and many a factor analyst was to make his reputation with a workable analytic scheme of factor rotation (p.26)".

many tests will not depend on all the factors. The factor matrix [$\mathbf{\Lambda}$] should have as many zero coefficients as possible. Such a matrix can be then considered as giving the simplest structure and presumably the one with most meaningful psychological interpretation" (Jöreskog, 1966, p.167).

Jöreskog, who worked on target rotation for some time (Mulaik, 1986), made two crucial contributions that helped to establish the dominance of the confirmatory approaches in factor analysis: (a) when lacking concise hypothesis regarding which elements should be fixed, researchers should divide its sample into two halves. The first split should be used to conduct EFA as a hypothesis-generating mechanism. The second half should be used to test the hypothesis previously generated (Jöreskog and Lawley, 1968); and (b) that $\mathbf{\Lambda}$ elements to be fixed as zeroes should be selected based on researcher's hypothesis or theory of interest (hence, the name of "confirmatory" factor models). The substantive advances in maximum likelihood estimation techniques in this area facilitated that many researchers quickly adopted these alternative models (Anderson and Rubin, 1956; Howe, 1955; Jöreskog, 1977; Lawley, 1958). Over time, CFA models were deemed as superior to their exploratory counterparts as they allowed for inspection of model fit indexes, the application of modification indexes to free incorrectly fixed parameters (e.g., free non-diagonal elements of $\mathbf{\Psi}^2$) and the possibility of conducting parameter invariance studies (Asparouhov and Muthén, 2009; Marsh et al., 2014).

CFA allowed analysts to go beyond the simple structure concept and to approach a new type of factor solution: the simplest structure[14]. In the simplest structure case, each item presents a unique factor loading (i.e., each Λ row presents, at most, a single non-zero entry; Ertel, 2013; Jennrich, 2018). While some authors suggested caution against the unjustified use of the simplest structure[15], these had little repercussion in the literature. Ultimately, the combination of confirmatory models following a simplest-structure approach to factor analysis dominated the field for decades (Asparouhov and Muthén, 2009; Marsh et al., 2014).

1.3.3 Simple Structure in Confirmatory Factor Models

The use of simplest structure in factor analysis has recently come to heavy scrutiny, and its use, emphatically discouraged (Asparouhov and Muthén, 2009; Guo et al., 2019; Marsh et al., 2010, 2014, 2009; Xiao et al., 2019). As many authors have argued, it could be foolish to

[14]The term simplest structure was renamed as "independent clusters structure" in the early 80's order to avoid confusion with Thurstone's original criteria (McDonald, 1984, p.82).

[15]"We either believe that small coefficients are zero in the population, or we do not. If we do, we should not get nonzero estimates of the zero coefficients. If we do not, we should not be using a simple structure" (McDonald, 1984, p.83).

assume, either by the effect of item content or direction[16] that items can act as pure indicators of a single common factor in a multidimensional structure (Asparouhov and Muthén, 2009). As argued in the literature (Jennrich, 2004b, 2006), such *ideal* situation, as denominated by some classical authors (Holzinger and Swineford, 1937; Howe, 1955; Jöreskog and Lawley, 1968), constitutes nothing but an illusion (Marsh et al., 2019; Morin et al., 2016).

As in any statistical procedure, there is no free-lunch in factor analysis. The enforcement of a simplest-structure on $\mathbf{\Lambda}$ has substantive detrimental consequences if the imposed restrictions are not consistent with the data: (a) CFA model fit is severely distorted; (b) the strategy of modifying a structure following modification indexes implies a research practice prone to capitalization on chance (Marsh et al., 2019, 2014); (c) bias is introduced not only in the unconstrained $\mathbf{\Lambda}$ parameters but also result in biased $\mathbf{\Phi}$ parameters. Biased $\mathbf{\Phi}$ estimation could lead to believe that higher-order structures are present when it is not the case or vice versa; and (d) in the case of CFA being integrated within an SEM model, biased estimation is translated into erroneous structural parameter estimation (Bandalos, 2018; Reise et al., 2018). Today, CFA is only considered to be preferable to EFA (due to its higher parsimony) when the simple structure holds at a population level and the model is correctly specified (Marsh et al., 2013). Unfortunately, this might never be the case.

In recent years, other types of exploratory or semi-exploratory techniques (i.e., Exploratory SEM, Bayesian SEM) have been suggested to be superior alternatives to the CFA-simplest structure-based approach (Guo et al., 2019; Marsh et al., 2019; Xiao et al., 2019). These techniques combine the benefits of CFA models (namely, fit inspection, invariance testing and inclusion of measurement models within structural models) without the consequences of imposing a simplest factor structure (Asparouhov and Muthén, 2009; Marsh et al., 2009). Either way, the paramount evidence discouraging the use of these types of confirmatory approaches in common factor research has ultimately invigorated a newly-found interest in EFA-based techniques.

1.4 The Case for the Exploratory Bi-factor Model

It is important to clarify that bi-factor solutions (i.e., Eq.1.3) represent factor structures where a single factor and several more-or-less simple group factors coexist. The general factor is the only factor with non-zero entries for all J items. Under confirmatory settings, items are expected to present substantive loadings in the general factor and a unique group factor (thus,

[16] For example, the use of mixtures of positive and negative items gives rise to load into unsubstantial method factors. Another strategy with similar consequences is using items with similar wording so to improve scale internal consistency.

items are of complexity 2), with the remaining non-vanishing entries being of exact value 0. However, and as Holzinger & Swineford discovered in their seminal paper (Holzinger and Swineford, 1937, p.53) such ideal bi-factor solution is often nowhere to be found (Reise et al., 2018, 2011), and was considered to represent a "hypothetical" case at best (Holzinger & Swineford, 1937; Table I). In this sense, Mss. Swineford reinforced this idea of non-simple bi-factor models as useful statistical models, particularly when compared Thurstone's simple correlated-factor model (Swineford, 1941).

Due to the increasing interest in bi-factor and exploratory models, several new approaches towards approximating bi-factor exploratory models have recently emerged (Jennrich and Bentler, 2011, 2012; Reise et al., 2011). All these initial perspectives presented three main shared characteristics: (a) they were aimed to overcome the limitations of the first approach developed (the Schmid-Leiman orthogonalization); (b) they relied on factor rotation as the mechanism to reveal full-rank bi-factor structures; and (c) they were developed under an umbrella of a minimal set of definitions, without an articulate theory substantiating what is a bi-factor model. These initial approaches to bi-factor exploratory are presented below.

1.4.1 The Schmid-Leiman Orthogonalization

The Schmid-Leiman method (i.e., SL; Schmid and Leiman, 1957) was the first approximation to conduct exploratory bi-factor analysis. An SL solution presents the form defined in Eq.1.3, but it is obtained by estimating successive higher-order models (i.e., oblique factor solutions) and a final transformation (i.e., the SL orthogonalization) form[17]. In detail, and starting from Eq.1.2, where $\mathbf{\Phi}_1$ refers to the factor-covariance matrix obtained from decomposing the observed correlation matrix, the first-order (initial) factor-correlation matrix can be decomposed as:

$$\mathbf{\Phi}_1 = \mathbf{\Lambda}_2 \mathbf{\Phi}_2 \mathbf{\Lambda}_2^T + \mathbf{\Psi}_2^2 \tag{1.7}$$

where $\mathbf{\Lambda}_2$ is the matrix of second-order factor loadings, $\mathbf{\Phi}_2$ is the second-order factor correlation matrix and $\mathbf{\Psi}_2^2$ is the matrix of second-order unique factors. If $\mathbf{\Phi}_2$ non-diagonal values are of relevance, this process could be repeated to obtain third-level factor loadings. However, in most applications, only a single second-order factor is extracted. Under this assumption, and that $\mathbf{\Phi}_2$ is a scalar, with λ_2 being a $P \times 1$ vector. Thus, $\mathbf{\Psi}_2^2$ can be expressed as:

[17] For interested researchers, an alternative procedure without requesting obtaining oblique solutions at each level is presented in Wherry (1959).

$$\boldsymbol{\Psi}_2 = [\boldsymbol{I} - Diag(\lambda_2 \lambda_2')]^{\frac{1}{2}} \tag{1.8}$$

Considering that $\lambda_G = \boldsymbol{\Lambda}\lambda_2$ (the vector of indirect effect of the second order factor onto the items), and $\boldsymbol{\Lambda}_{GRP} = \boldsymbol{\Lambda}\boldsymbol{\Psi}_2$ (representing the unique effect of first-order factors onto the items) can expressed in block form such as:

$$\boldsymbol{\Lambda}_{SL} = \begin{bmatrix} \lambda_G & \boldsymbol{\Lambda}_{GRP} \end{bmatrix} \tag{1.9}$$

which is in similar form of Eq.1.4. Therefore, the reproduced correlation matrix can be expressed as:

$$\boldsymbol{R} = \boldsymbol{\Lambda}_{SL}\boldsymbol{\Lambda}_{SL}^T + \boldsymbol{\Psi}^2 \tag{1.10}$$

More importantly, and as explained in detail in (Mansolf and Reise, 2016; Yung et al., 1999), the SL transformation expects a simple structure at each level (Schmid and Leiman, 1957, p.54)[18]. This imposition is relevant when understanding the implicit relationship between λ_G and $\boldsymbol{\Lambda}_{GRP}$ estimated via an SL orthogonalization: for items loading in the same factor in $\boldsymbol{\Lambda}$, their factor loadings in λ_G and $\boldsymbol{\Lambda}_{GRP}$ are proportional. This effect, which is the result of using $\boldsymbol{\Lambda}$ in λ_G and $\boldsymbol{\Lambda}_{GRP}$ computation, results in a series of linear constraints being observed in the $\boldsymbol{\Lambda}_{SL}$ solution: the *proportionality constraints*.

The proportionality constraints have been substantially discussed in the literature as (a) any SL-based solution encompasses them, despite not always been directly observable (Mansolf and Reise, 2016); (b) their presence identifies the generative mechanism from which the SL solution is obtained, namely the SL orthogonalization of a higher-order solution into a bi-factor form; and (c) their presence reveals that SL solution is a restricted version of a bi-factor model where the proportionality constraints are not imposed (Yung et al., 1999)[19], where $rank(\boldsymbol{\Lambda}_S L) = rank(\boldsymbol{\Lambda}) = rank(\boldsymbol{\Lambda}_{BF} - 1)$. Ultimately, $\boldsymbol{\Lambda}_{SL}$ parameters will always represent biased estimates of any full-rank bi-factor model. As such, its overall usefulness has been traditionally questioned (Jennrich and Bentler, 2011, 2012; Mansolf and Reise, 2016; Reise et al., 2018, 2010, 2011; Yung et al., 1999). However, as the SL transformation is not unique (i.e., $\boldsymbol{\Lambda}$ must be rotated in the first-order solution estimation step), this indeterminacy could be exploited to obtain an SL-solution that represents the close approximation to an unconstrained bi-factor model (in the least-square sense; Waller, 2017;

[18] Even though a "simple structure" is never mentioned in the original article, its procedure implicitly assumes non-overlapping sets of loadings for the group factors.

[19] As demonstrated by the use of the *generalized* SL transformation proposed by Yung et al. (1999), an SL solution is nested within an unrestricted second-order model which, at the same time, is equivalent to the bi-factor model (Chen and Zhang, 2018).

Giordano & Waller, 2019). Thus, the debate of whether the SL orthogonalization could approximate unrestricted bi-factor solutions with a certain degree of accuracy is still an open question in the literature.

1.4.2 A Brief note on Rank Deficiencies and Bi-factor Models

Before introducing current methods for estimating exploratory bi-factor methods, the distinction between restricted and unrestricted bi-factor models should be briefly discussed. Firstly, structures presenting simultaneous direct effects from the general factor on both, the items and the group factors, are not included considered here (e.g., Structure C, Figure 1, Yung, Thissen and McLeod, 1999), as additional restrictions should be imposed for this model to be identified (Eid et al., 2018; Markon, 2019). Such structures will not be discussed in this thesis dissertation. Thus, for a factor model to be in bi-factor form, it should follow the definitions provided in Eq.1.4 and Eq.1.5, regardless of the potential constraints present between sets of factor loadings (Giordano and Waller, 2019; Waller, 2017).

Bi-factor models are often classified based on of their rank, distinguishing between full (unrestricted) and rank-deficient (restricted) bi-factor models (Giordano and Waller, 2019; Waller, 2017). Such distinction is of relevance as $\mathbf{\Lambda}_{SL}$ will always be, by definition, a rank-deficient matrix. Indeed, $\mathbf{\Lambda}_{SL}$ will always yields a solution with the same rank of $\mathbf{\Lambda}$, not $\mathbf{\Lambda}_{BF}$ (Waller, 2017), as λ_G is indeed spanned in the space-column of $\mathbf{\Lambda}$. Either way, it should be clarified here that rank-deficiency is only a necessary, but not a sufficient condition, for a given $\mathbf{\Lambda}$ in bi-factor form to be consistent with having being obtained via SL transformation. The set of deficient-rank solutions allows for solutions consistent and inconsistent with having being generated from a higher-order model via an SL transformation: If a bi-factor factor loading matrix is rank-deficient, with λ_G being derived from $\mathbf{\Lambda}_{GRP}$), and the coefficients of the linear combination being derived from λ_2 and $\mathbf{\Psi}_2$, then this bi-factor factor matrix is consistent with an SL transformation. But any given $\mathbf{\Lambda}_{BF}$ can be rank deficient as the result of other circumstances, as alternative sets of coefficients to the so-called proportionality constraints can be applied to derive $\mathbf{\Lambda}_G$. Moreover, $\mathbf{\Lambda}_{BF}$ could be rank-deficient in terms of two vectors of $\mathbf{\Lambda}_{GRP}$ being collinear to each other. Rank computation has its limitations whenever distinguishing if a bi-factor form $\mathbf{\Lambda}$ is consistent to have been obtained from an SL transformation or not. Even though this situation could be a mathematical curiosity with limited impact in applied settings, as it is challenging to imagine a generative process underlying different rank-deficient bi-factor models outside the SL transformation.

1.5 New Methods to approximate Exploratory Bi-factor Model

Two separate schools of thought intended to substitute the SL transformation as the main approximation to exploratory bi-factor models. Both strategies revolved towards the use of bi-factor form rotations: one through designing analytical bi-factor rotation criteria and the other by applying bi-factor partially specified target rotation.

1.5.1 The bi-factor criteria

In two seminal articles, Jennrich and Bentler (2011, 2012) established the foundations of bi-factor exploratory modelling. In their research, they adapted two existent well-known criteria to the orthogonal and oblique bi-factor cases. These criteria were defined as:

$$B(\mathbf{\Lambda}_{BF}) = Q(\mathbf{\Lambda}_{GRP}) \tag{1.11}$$

where Q stands for a given rotation criterion. For dealing with simple and orthogonal structures, Jennrich and Bentler adapted the quartimin criteria to the bi-factor case:

$$B(\mathbf{\Lambda}_{BF}) = Q(\mathbf{\Lambda}_{GRP}) = \sum_{j=1}^{J} \sum_{p=1}^{P} \sum_{p'=p+1}^{P} = \lambda_{jp}^{2} \lambda_{jp}^{\prime 2} \tag{1.12}$$

and called it bi-quartimin criterion[20]. With the same spirit than the original criterion, bi-quartimin is maximized when items have, at most, complexity two (Jennrich and Bentler, 2012). To recover structures presenting items of higher complexity, the authors developed another widely known criterion, Yates' geomin (Yates, 1987) as:

$$B(\mathbf{\Lambda}_{BF}) = Q(\mathbf{\Lambda}_{GRP}) = \sum_{j=1}^{J} \prod_{p=1}^{P} (\lambda_{jp}^{2} + \varepsilon)^{\frac{1}{P}} \tag{1.13}$$

where ε corresponds to a rotation fixed arbitrary constant of small value (i.e., .01) which allows the function to be differentiable. As occurred with bi-quartimin, this criterion was denominated as bi-geomin. Bi-geomin and bi-quartimin quickly gained popularity, as both rotations were easy to use and set as default rotation when conducting both, exploratory bi-factor analysis and ESEM in *Mplus* (Asparouhov and Muthén, 2009).

[20]Not to be confused with Carroll's bi-quartimin criterion (Carroll, 1957).

1.5.2 The bi-factor specified target rotation

One of the first known attempts to overcome the SL orthogonalization was suggested by Reise et al. (2010). These authors realized that whereas rank-deficient solutions represented biased estimations of the true full-rank bi-factor parameters, they were useful to identify the overall pattern of negligible and vanishing entries of $\mathbf{\Lambda}$. This idea, further extended by Reise et al. (2011), suggested using the SL to define a "Procrustes" o target rotation in bi-factor form that would allow estimating a full-rank bi-factor solution. The target rotation (Browne, 1972, 2001; Cureton and Mulaik, 1971) is a widely-known rotation procedure that enables researchers to incorporate previous information regarding the expected form of the factor loading matrix into the rotation process. Accordingly, the target rotation has been considered as a "semi-confirmatory" rotation (Browne, 2001).

The target rotation traditionally aims to provide the best least-squares approximation of $\mathbf{\Lambda}$ to a user-defined target matrix $\mathbf{B}$ as follows:

$$f(\boldsymbol{B}) = \sum_{j=1}^{J} \sum_{p=1}^{P} w_{jp}(\lambda_{jp} - b_{jp})^2 \tag{1.14}$$

where b_{jp} represents a given entry of the target matrix $\mathbf{B}$ and w is an operator that takes value 1 if the correspondent b element is defined in $\mathbf{B}$ and 0 otherwise (Browne, 2001). $\mathbf{B}$ elements are given a "target" value towards which $\mathbf{\Lambda}$ entries are orthogonally or obliquely projected [21] The most common values used for specifying $\mathbf{B}$ are one and zero, correspondent to $\mathbf{\Lambda}$ elements to be completely maximized or minimized, respectively. Nevertheless, if (Browne, 1972) algorithm is applied, not all $\mathbf{B}$ entries are requested to be defined: One $\mathbf{B}$ matrix only specifying a partial set of elements to be minimized or maximized can be used as target matrix[22]. To decide which entries should be minimized or maximized, a single cut-off for distinguishing vanishing and relevant factor loadings was often settled (i.e., .30 as in McDonald, 1999)[23]. Entries who are lower in value were traditionally fixed to zero, while entries above it are either fixed to one or freed (not given any target value) depending on researcher's preference for a partial or a completely specified target rotation (Browne, 2001).

[21] The name Procrustes rotation comes from the Greek myth of the same name. Procrustes was a bandit that, after housing travellers, terrorized them by stretching or amputating their legs so the unfortunate visitor would fit the iron bed in which they were offered to sleep. Despite Procrustes efforts, and as it often occurs with the rotation named after him, nobody ever fitted the bed perfectly.

[22] The partially specified target matrix was, for a time, called the Xerxes method, as suggested by Prof Raymond B. Catell (Derflinger and Kaiser, 1989). Prof Cattel was inspired by the charts showing the disposition of the Persian fleet in the Battle of Salamis. Unfortunately, this tradition has been lost in the literature.

[23] .30 was proposed as a cut-off because as factor loadings represent the correlation coefficients between items and factors, a .30 factor loading would imply that the factor explains almost 10% of the item variance (Bandalos, 2018).

Reise et al. (2011) suggested using a partially specified target rotation based on an initial SL solution as a valid method for approximating full-rank bi-factor exploratory models. In this sense, $\mathbf{B}_{BF}$ is also constructed in block form such as:

$$\boldsymbol{B}_{BF} = \begin{bmatrix} \boldsymbol{B}_G & \boldsymbol{B}_{GRP} \end{bmatrix} \tag{1.15}$$

where $\mathbf{B}_G$ elements are always freed (as all items are expected to load into the general factor), and $\mathbf{B}_{GRP}$ elements are defined based on an SL solution and a single cut-off point (e.g., .15). Thus, a rotation towards $\mathbf{B}_{BF}$ is expected to recover a full-rank solution accommodating a single, general factor plus a simple structure for the group factor loadings. This procedure was demonstrated to improve factor loading recovery under many circumstances when compared with an SL solution (Reise et al., 2011).

1.6 Challenges on Exploratory Bi-factor Modeling

As the interest in exploratory raised, the geomin and target criteria started to be more scrutinized in the literature (Asparouhov and Muthén, 2009; Hattori et al., 2017; Moore et al., 2015; Myers et al., 2013, 2015). As such, it was only matter of time until the first studies investigating the performance of the available methods for BEFA (i.e., SL, bi-geomin, bi-quartimin, SL with target rotation) started to appear in the literature Mansolf and Reise (2016).

1.6.1 The Fall of the Bi-factor Rotation Criteria

As acknowledged by Jennrich and Bentler, bi-geomin should be preferred to bi-quartimin, as the simplest bi-factor structure is expected to be violated (Jennrich and Bentler, 2012). Due to bi-geomin and geomin popularity (Marsh et al., 2014; Morin et al., 2016), both rotation have been deeply scrutinized in the literature: (a) geomin is considered to fail to recover structures with three or more factors and items of complexity three or higher (Asparouhov and Muthén, 2009); (b) geomin performance is strongly dependent upon ε choice, with suggestions of optimal ranging from .50 to .001 values (Hattori et al., 2017). For example, ε is chosen in *Mplus* depending upon the number of columns (Asparouhov & Muthén, 2009, p.409, footnote 6) up to .01 with four or more factors. On the contrary, Browne (2001) suggest ε to be .01, and to be increased slightly for more than three or four factors. Furthermore, other authors have found that geomin with a higher ε value (i.e., .50) could outperform traditional geomin in most settings (Celimli Alkoy, 2017; Marsh et al., 2010, 2009); and (c) geomin and bi-geomin were considered as highly dependent on starting values and prone to result in local minima solutions (Hattori et al., 2017; Mansolf and Reise, 2016). Whereas local minima could be seen as an opportunity rather than a disadvantage in factor analysis (Rozeboom, 1992), it hinders the application of these methods in applied contexts.

Due to their relevance, bi-geomin local minima issues will be hereafter explained in greater detail. It is noteworthy that bi-geomin (and bi-quartimin) are minimized solely based on the simplicity of $\mathbf{\Lambda}_{GRP}$ (see Eq.12 and Eq.13). As explained in detail in Mansolf and Reise (2016) or Robertson (2019), the gradient projection algorithm plays a relevant role in the presence of local minima and collapsed solutions in both rotation criteria. This optimizer is based on two iterative steps (Jennrich, 2001, 2002, 2004a; Mulaik, 2010), namely a minimization and a projection step. During the former, the criterion is minimized for $Q(\mathbf{\Lambda}_{GRP})$ using a gradient descent strategy. The latter is crucial, as the resulting transformation matrix **T** found in the first step is not expected to be in the admissible set of solutions (as $\mathbf{TT'} \neq \mathbf{I}$ in the orthogonal rotation case). To ensure its appropriateness, during the projection step the complete **T** (including not only $\mathbf{\Lambda}_{GRP}$ but $\mathbf{\Lambda}_G$ columns) is projected via a Procrustes rotation into the non-linear manifold of potential admissible solutions. As a consequence, variance might shift from $\mathbf{\Lambda}_{GRP}$ to the $\mathbf{\Lambda}_G$ during this step. As such, λ_G is "implicitly" rotated when applying bi-geomin and bi-quartimin criteria. Noteworthy, bi-quartimin does not seem to present local minima issues as stringent as bi-geomin (Mansolf and Reise, 2016; Weide and Beauducel, 2019), suggesting that the geometry of the criterion itself also plays a role in the presence of these effects (Hattori et al., 2017).

Lastly, it should be noticed that bi-geomin and bi-quartimin will always result in full-rank solutions. Under the assumption of the true population model being a higher-order model transformed via an SL (i.e., where proportionality constraints held true), their solutions can be either not identified (for example, when $rank(\mathbf{\Lambda}_{GRP}) = 2$; Jennrich & Bentler, 2012) or prone to present Heywood cases due to over-extraction issues (Mansolf and Reise, 2016) These concerns have led to many researchers to suggest that practitioners should favour the use of partially specified target rotation (Guo et al., 2019; Marsh et al., 2019).

1.6.2 Questions in Target Rotation

The bi-factor target rotation presented in Reise et al. (2011) offered a suitable alternative to the bi-factor rotation criteria: (a) it did not pose the technical challenges associated with bi-geomin and bi-quartimin "implicit" rotation of the general factor; (b) it allowed the incorporation of previous knowledge in the bi-factor rotation procedure (Browne, 2001); and (c) it was available in most factor analysis-related software, as the rotated solution can be obtained based on a singular value decomposition (Cliff, 1966; Schönemann, 1966; ten Berge, 2006; ten Berge and Nevels, 1977). On the other hand, the adequacy of the target rotation has been strongly questioned, as its performance depends on the appropriateness the target matrix defined by the researcher (i.e., the number of location of targets given, the value of the targets, etc.). The target rotation has been a polarizing method throughout the literature (Mulaik, 1986). Many scholars have argued that the target rotation could result in factor patterns concordant with researchers' expectations regardless of the true nature of the true item-factor relationships (Harman, 1967; Hurley and Cattell, 1962; Moore et al., 2015; Mulaik, 2010).[24].

On the other hand, other scholars interpret this flexibility as a positive feature, depicting target rotation as a mechanism for a healthy reintroduction of researchers' expertise into factor rotation (alas early-days graphical rotation) after 50 years of abusing of mechanical, automatic rotations (Mulaik, 2018) or Gorsuch (1983, p.245). Nevertheless, it should be clarified that target rotation does not constitute a singular, monolithic rotation procedure. It should be rather understood as a family of related methods sharing a minimization criterion (resembling similarities to what many authors consider to the geomin case and ε values). In this sense, there exist many types of target rotations, each defined based on the types and number of targets applied, and the method applied for finding such targets.

Either way, it was not until recently that the first simulation studies investigating the consequences of these decisions on target rotation appeared in the literature. These studies have allowed researchers not only to have a better knowledge of the statistical behaviour of the target rotation but to expand target rotation with new capabilities to resolve some of its drawbacks.

Defining Target Rotations

Target rotation has an ample history in psychometrics (Horst, 1941; Jennrich, 2018; Lawley and Maxwell, 1964; Tucker, 1940), playing a crucial role in the development of many current rotations (Browne, 2001; Fleming, 2012) and in the emergence of confirmatory factor analysis (Mulaik, 1986). In this sense, the target rotation was one of the first rotation methods available (Mosier, 1939) and it has been studied in detail throughout the factor analysis literature (Horst, 1941; Korth and Tucker, 1976; Tucker, 1944)[25]. While early rotations were based on completely specified solutions (rotations where each loading is given a target value),

[24]Noteworthy, Prof John R. Hurley and Prof Raymond B. Cattell abstained themselves from publishing their software to compute Procrustes rotation for years to avoid its spread and misuse by other researchers. However, the excessive number of "quick journal publications, apparently unhindered by editors, of orthogonal simple structure solutions by Kaiser's Varimax" (p.260) changed their mind (Hurley and Cattell, 1962).

[25]As detailed in Jennrich (2007), Mosier's method provided the best approximation of $\mathbf{\Lambda}$ to $\mathbf{B}$ not in the least-squares sense, but in terms of maximizing factor congruence (ten Berge and Nevels, 1977).

since the early '70s, the partially specified target rotation methods have being preferred (Browne, 1972; Meredith, 1977).

Through its history, the target rotation has been tied, even more explicitly than any other rotation, to the concept of simplicity (Jennrich, 2004b, 2006). For clarity's sake, simplicity and complexity have been using indistinguishably throughout the literature. For example, a row presenting a single non-vanishing value could be understood to be of simplicity = 1 or complexity = 1. Both mean that a single non-vanishing entry was found in that specific row. Simplicity has been characterized in terms of item or row-simplicity, defined as the is the count of existent non-vanishing entries per row or in terms of the column or overall pattern matrix simplicity.

Item simplicity has been a traditional workhorse of factor rotations (see the fifth rule)[26]. Be that as it may, maximize row simplicity is only akin to approximating a simple structure in the simplest structures case (Fleming, 2012). Under complex structures including several non-vanishing entries per row (as the bi-factor model), one could expect that a few loadings to be of high magnitude, but several to be in the mid-range (i.e., cross-loadings). Under these conditions, it is unclear whether such a structure should be considered as simple or not (Jennrich, 2004b, 2006). Under realistic conditions, maximizing item simplicity might not be equivalent to seek for the simplest structure. Lastly, and as reflected by McDonald (1984): "the simply structure concept has no implication at all for the sizes of elements that are thought to be nonzero" (p.84). Accordingly, several scholars have argued that factor rotation objective should be to find as many zeros as possible (Fleming, 2003, 2012). As best explained by Jennrich (Jennrich, 2004b):

> "Perfect simple structure and Thurstone simple structure don't occur in practice. They are at best idealizations. Unfortunately, there is not generally accepted broadly applicable definition of simple structure. It is generally felt, however, that a loading matrix with many small values and a small number of larger values is simpler that one with mostly intermediate values. Motivated by this, we will consider methods that produce small or large loading and hopefully not too many intermediate loadings" (p.264).

The idea of maximizing the number of zeroes can be traced back to Thurstone's ideas on how to approximate the simple structure concept: "A simple structure can be defined

[26] Kaiser (1958), following Guildford's criticism of the simple structure concept, developed Varimax not to approach a simple structure, but factor invariance. As commented in the abstract of his seminal paper: "It is proposed that the ultimate criterion of a rotational procedure is factorial invariance, not simple structure". Noteworthy, it is clear to this thesis author that such was Thurstone's original goal when proposing the idea of simple structure

statistically as a factor matrix in which a very large number of zeroes entries appear" (Thurstone, 1940, p.195). In factor rotation, this problem was posed as the maximization of the hyperplane count (Catell, 1966; Eber, 1966; Gorsuch, 1983). Hyperplanes represent the intersection of all factors with a given factor within a structure (Gorsuch, 1983). Items with near-zero loadings in such hyperplane would present a high number of near-zero entries in several $\mathbf{\Lambda}$ columns, and vice versa. Hyperplane counts are commonly operationalized as the number of near-zero loadings. Similar to factorial simplicity, it represents a statistic that could be computed per row, column or for the total number of entries in $\mathbf{\Lambda}$. Noteworthy, the latter case has been the most common conceptualization of hyperplane count in the literature (Fleming, 2012; Rozeboom, 1991). Indeed, Thurstone aimed to maximize the hyperplane count on reference vectors as an indirect manner of finding items with a high number of zero loadings (Gorsuch, 1983, p.237). Again, since Jennrich and Sampson (1966) innovations, hyperplane counts were aimed to be estimated directly in $\mathbf{\Lambda}$ instead. Unfortunately, the direct maximization of hyperplane counts could result in suboptimal performance if the location of the zeroes is not taken into account (Eber, 1966; Fleming, 2003). In the end, most methods for estimating hyperplane counts are also affected by the magnitude and location of the non-zero elements (McDonald, 1984).

It is in this context, the target rotation provided a compromise between finding an optimal number of near-zero elements while taking into account the location of non-zero loadings. Given its semi-confirmatory nature, the target rotation allows researchers to specify which $\mathbf{\Lambda}$ elements should close to zero (indirectly maximizing the hyperplane count) after the rotation process (Catell, 1966; Eber, 1966). The remaining question was, precisely, how to decide which loadings should be counted as zero and which should not in the target matrix. As better reflected by Hendrickson and White (1964): "The trouble seemed to lie in the definition of "salient" [loadings] (p.66)".

How to define an optimal target rotation has been a recurrent topic of research in factor rotation. For a long time, factor analysts have largely rejected the use of universal, theory-based thresholds for separating relevant and near-zero loadings[27] (Cudeck and O'Dell, 1994; Izquierdo et al., 2014), as factor loading magnitude are dependent upon structure characteristics such as the number of factors extracted, sample size, etc. Nevertheless, a mid-range value can only be understood in a relative sense (in the context of the range of loadings for a given factor). In this sense, using a single, fixed cut-off point could result in too liberal or too stringent cut-offs if factor loadings differ in their average factor loading (which is often the case). The principal solution to this problem has been defining an "empirical"

[27]Initial considerations in hyperplane counts considered using $\pm$ .10 as the reference cut-off, as in Catell (1966).

factor loading salience cut-off based on the estimated factor solution. The best exponent of this approach might be Promax (Hendrickson and White, 1964). Promax starts by finding an initial orthogonal rotated solution using Varimax (or any other rotation procedure). This initial solution is raised to a certain power (i.e., 4) leaving signs unchanged, before constructing a fully-specified oblique target matrix based on this solution. The combination of an orthogonal solution to an oblique rotation caused these methods being called orthoblique in the literature. Since its proposal, Promax has become on the most applied rotation methods, and its found in mainstream statistical software (i.e., SPSS, R, etc.). However, Promax presents some technical limitations, such as applying Mosier (1939) rather than Browne (2001) method (Lorenzo-seva, 1999), and the issue on how to decide which power should be applied.

A third (and a major) breakthrough in this context was the apparition of the Simplimax rotation (Kiers, 1994). Simplimax aimed to recover the simplest partially specified target matrix and the optimal target rotation without depending upon an initial solution. Simplimax provides the "best" simple target matrix for given hyperplane count (Kiers, 1994, p.568) by minimizing the sum of squares of the H smallest loadings in the solution. Thus, Simplimiax only requests the optimal hyperplane counts, while the position of these zeroes in the target are found via a minimization[28]. Therefore, the only question remaining was how to how to find the hyperplane count (i.e., H). Kiers suggested using a scree-test to compare the minimized rotation function values at each different H count, and to decide where a noticeable increase of the function value might occur. Simplimax has repeatedly demonstrated to be one of the most flexible rotation methods available to model complex structures (Browne, 2001; Fleming, 2012; Jennrich, 2004b, 2006).

The same year that Simplimax was published in Psychometrika, Trendafilov (1994) suggested an alternative method to Promax. In this case, the proposal was to substitute the approximate zeroes used in a target matrix obtained using Promax by exact zeroes via a vector majorization (Promaj; Trendafilov, 1994); The vector majorization allowed to identify which loadings should be fixed to zero in the target without requesting the use of any power. Additionally, vector majorization introduced computing a different cut-off for defining which loadings should be fixed to zero not only empirically, but for each factor in $\mathbf{\Lambda}$ separately (a similar idea was suggested by (Derflinger and Kaiser, 1989). This idea was incorporated in the Promin rotation (Lorenzo-seva, 1999; Lorenzo-Seva and Ferrando, 2019a). Promin aims to find a column-specific cut-off point to decide which items should be fixed to zero in an oblique partially specified target. Specifically, each cut-off is set to a quarter of a standard deviation above the mean of factor loadings. Noteworthy, Promin has achieved

[28]Kiers proposed a second version of Simplimax in which researchers could specify the expected row complexity (Kiers, 1994, pp.576-577).

fantastic results when compared with other rotations, and its use is strongly recommended here (Fleming, 2012). Other relevant mentions of hypercount-based rotations are the Horst-Harris rotation (Derflinger and Kaiser, 1989), the Hypermax (Fleming, 2012) or the Hyball rotation (Rozeboom, 1991).

Unfortunately, there are several concerns regarding these approaches, particularly with regards to how the partially specified target matrices are defined. For example, all reviewed rotations relied on some arbitrary parameter chosen on authors' discretion (namely the power in Promax or the standard deviation modifier applied in Promin). Secondly, these algorithms did not explicitly account for the trade-offs between factor simplicity and target misspecification, where the consequences of fixing a relevant or freeing a near-zero element should be considered with greater detail. In this sense, most methods are based on conservative strategies. Even though one might argue that this approach might be suboptimal: In the case of fixing a relevant cross-loading to zero, the least-square difference between the true value and the wrong zero value will be translated to the computation of the rotation criterion. Moreover, the resulted factor loading (after the rotation) would be a biased estimation of its true value. In the case of mistakenly freeing a near-zero value, the impact of this error would be rather limited, as its value will not be taken into the criterion computation, and the error between the true zero value and the rotated loading will be small anyway. If enough correctly fixed values are present in the structure, its value would remain close to zero in either case (as rotational freedom exhausts). Strategies based on freeing major cross-loadings while fixing minor cross-loadings are expected to be adequate when recovering non-bi-factor ESEM models (Guo et al., 2019). Accordingly, any partially specified target rotation based on an empirical cut-off should be primarily concerned with avoiding fixing relevant cross-loadings to zero, rather than ensuring that all near-zero values are correctly fixed to zero.

The flexibility of the target rotation for approximating values to their target values (as in Promax) has been recently exploited in the context of the partially specified target rotation as to find simpler solutions (Hendrickson and White, 1964). This idea was developed in detail by Moore et al. (2015) and Moore (2013) in a process called Iterative Target Rotation: based on a rational start (e.g., a rotated solution), an initial target matrix is built based on a single salience fixed cut-off point so to locate near-zero values in the matrix[29]. Secondly, a first partially specified target rotation is performed. Next, the rotated pattern is again compared with the fixed cut-off point to define a new partially specified target matrix, and to perform a second target rotation. This strategy, which is repeated until the solution converges (i.e., two consecutive rotated solutions being equal), results in initial mid-range values being

[29]Interestingly, in the Appendix of his PhD thesis, Dr Moore developed an alternative empirical criterion for defining which loadings should be fixed based on factor loadings confidence intervals. To our knowledge, this idea never came into fruition in an official publication, but it could be interesting to explore in future research.

minimized, and to outperform other rotations under complex situations (Moore, 2013; Moore et al., 2015). Additionally, the authors demonstrated this strategy converged in a stable solution in less than seven iterations in all cases studied, regardless of use the use a rational start. Moore et al. (2015) illustrated that refining partially rotated target-based solutions, as expressed not only by Browne (2001), as often recurred, but by many authors throughout the literature, could serve as a valid scheme for finding simple solutions in many instances.

Completely vs Partially Specified Target Rotation

A researcher interested in applying target rotation is faced with two choices: Firstly, whether to apply a completely vs a partially specified target rotation (Browne, 2001). Secondly, the values to be used for the targeted entries of the matrix. Unfortunately, both decisions have been often diminished to decide between a maximalist completely specified target rotation (whose targets are either of value one for factor loadings to be maximized or zeroes for loadings to be minimized) or a partially specified target rotation (where only the former restriction of zero targets is set). Be that as it may, the reality is that a researcher could decide to use a completely or partially specified target rotation whose values were different from zero or one. For example, Guo et al. (2019) showed that changing target values for near-zero loadings from 0 to .10 improved the estimation of ESEM models under certain circumstances. The possibilities of these types of target rotations have not been considered in the context of bi-factor modelling and are not detailed here.

At first, there is little information regarding whether to prefer the completely or partially specified versions target rotation, whereas the use of the latter has strongly preferred throughout the history of psychometrics. Nevertheless, the application of completely specified target rotation has several advantages: (a) the target rotation has a closed-form solution based on singular value decomposition (Browne, 1972; ten Berge, 2006), avoiding the use of gradient projection algorithm and its dependency on starting values, the descent step size and convergence rates (Browne, 1972; Cliff, 1966; Schönemann, 1966); (b) the conditions for ensuring identifications are already well-known (Gower and Dijksterhuis, 2004; ten Berge, 2006); and (c) there are several methods available for performing the rotation under several settings, as these methods are often applied in other sciences such as robotics and artificial vision (Casper and Gower, 2010; Gower and Dijksterhuis, 2004).

Additionally, providing as many elements as possible in the target matrix has been associated with improved rotation performance (Celimli Alkoy, 2017; Myers et al., 2013, 2014). In detail, as seen in Eq.1.14, the target rotation criterion is defined by **W** and **B**. **W** is crucial with regards to to criterion computation, as any b_{jp} elements corresponding to $w_{jp} = 0$ will not be accounted for in Eq.1.14. Increasing the number of non-missing elements (increasing the effective number of elements used to compute the criterion) has been shown to increase its accuracy, particularly under low communality conditions (Myers et al., 2013). Recent evidence shows that with sufficient correct fixed values, target rotation outperforms geomin with regards to accuracy. To be noticed that the former outperforms the latter with regards to estimation stability (Celimli Alkoy, 2017; Myers et al., 2015). Lastly, target rotation has been depicted to be resilient to target error (i.e., an error is defined by the difference between true $\lambda_{jp} - b_{jp}$). This robustness is associated with the flexibility of the least-squares criterion (Myers et al., 2016, p.503). Thus, increasing the information regarding the factor structure is more relevant than the specific values specified in the target rotation (Celimli Alkoy, 2017; Myers et al., 2015).

It could be reasonable to assume that a completely specified target rotation would result in improved performance than any partial target rotation alternative, as it will always convey more information in **B**. Unfortunately, whether this is true is yet unknown, as Myers et al. (2017, 2014) studies presented some relevant limitations. For example, they never included a completely specified target rotation, whereas being commonly applied by researchers. Moreover, this type of target rotation has been recently reintroduced in the bi-factor literature in the form of the Direct Schmid-Leiman and the Direct Bi-factor methods (Giordano and Waller, 2019; Waller, 2018). In these methods, the rank-deficient bi-factor model which is closer to a full-rank deficient method (in the least-squares sense) is found as follows: firstly, a completely specified target matrix is found employing a correlated-factor solution and a fixed cut-off point (.25) scheme; secondly, this rotation matrix is augmented by appending a column of zeroes; thirdly, a singular value decomposition-based target rotation is performed using this rotation matrix and an unrotated solution of the same expected dimensionality of the final bi-factor model. Either way, the consequences of applying these schemes of completely specified target rotation are unknown in the literature (see Chapter 5).

The Bi-factor Case

The partially specified target solution has emerged as a compelling method for conducting BEFA. However, previous target-based BEFA strategies paid little attention to the particular characteristics of this model. For example, the problematic simultaneous presence of a strong general and several less relevant group factors within a single factor structure (Rodriguez et al., 2015, 2016). Additionally, most methods did not explicitly accounted for the unexpected item complexities of items (i.e., the presence of cross-loadings or pure indicators). While minor cross-loadings should be expected to occur as in any other factor solution (Marsh et al., 2014, 2009), the presence of pure indicators is a unique feature of the bi-factor model.

A pure indicator is defined as an item with complexity one, being the non-vanishing factor loading located in the general factor loading vector (Chapter 2, 3 and 4).

Pure indicators are commonly found in the bi-factor literature, with examples of these types of disturbances being found in the Spearman's Visual Perception Test (Holzinger & Swineford, p.53), the Twenty Four Psychological Test (Harman, 1967; Jennrich and Bentler, 2011) the Quality of Life Dataset (Abad et al., 2017; Chen et al., 2006), and the Observer Alexithymia Scale (Jennrich and Bentler, 2012; Reise et al., 2010), among many others. There are many reasons to believe that pure indicators should be expected to occur in bi-factor modelling. Firstly, most tests are designed to be either unidimensional or multidimensional scales. Understandably, items are designed to reflect a general or a particular aspect of psychological phenomena (not both, as would be desirable in a bi-factor structure). Finding a set of items which would simultaneously act as markers of a general plus a single group factor could be quite a challenging task. Secondly, group factors often represent unreliable, minor sources of variance (Markon, 2019; Rodriguez et al., 2016), which are inherently harder to estimate than the general factor. Group factors not only account for a residual variance to the general factor but are also often defined by only a few mid-range factor loadings (Constantinou and Fonagy, 2019; Markon, 2019; Rodriguez et al., 2016).

Be that as it may, the recovery of true pure indicators is also relevant from another point of view: bi-factor structures presenting true pure indicators are ensured to be full-rank. Under the presence of a zero group factor loading, the general factor can no longer be obtained to a linear combination of the remaining group factor loadings for that group[30]. Thus, structures presenting pure indicators are inconsistent with having been generated via a higher-order model and SL orthogonalization. Therefore, there is an impending necessity of updating and adapting BEFA target rotation methods to consider the particular features of the bi-factor model and to substitute the use of fixed cut-off by an empirical cut-off point. The design of this empirical cut-off point should be guided by the knowledge of the costs associated with each type of error that could occur in target misspecification. Accordingly, liberal, smaller empirical cut-off points (where less meaningful cross-loadings are fixed to zero) are expected to provide good recovery of complex structures such as the bi-factor model.

A useful idea to understand how to find a suitable cut-off point might be the examination of Sorted Absolute Loading (i.e., SAL) plots to guide decisions in factor rotation (Jennrich, 2007). A SAL plot is a scatterplot revealing the empirical distribution of absolute factor loadings sorted by magnitude. These plots, similar to Simplimax scree plots for function values, could help in identifying jumps between vanishing and non-vanishing relevant cross-loadings. Additionally, and as proposed in Promaj or Promin, empirical cut-off points should

[30]The mathematical proof is omitted here.

be defined for each group factor separately. As the bi-factor model is likely to present group factors differing in their average factor loading, this improvement should be considered not as a suggestion, but as an obligation for any bi-factor target-rotation-based method. Furthermore, the inclusion of refinement schemes via iterative target rotation as suggested in Moore et al. (2015) or Moore (2013) should be strongly encouraged.

The last issue regarding the application of partially specified target rotation is that a researcher must understand if a particular rotated solution is identified or not. As said before, for the general (non-bi-factor) case, Anderson and Rubin (1956) rule 1 for identification establish that, at least, $\frac{P-1}{2}$ entries must be specified in each $\mathbf{B}$ column for orthogonal rotation. Additionally, the rank of $\mathbf{\Lambda}_P$, defined as the matrix retaining rows whose entries have been fixed to zero in a column P with these zeroes deleted, must be of value $P-1$ for all $P = 1,...,P$ (Asparouhov and Muthén, 2009). Additionally, other minor identifiability conditions are complemented by Peeters (2012). These identification conditions are also pertinent in target rotation, where instead of requesting zero values in $\mathbf{\Lambda}$, a similar number of fixed values is requested in $\mathbf{B}$ (for a detailed example, see Asparouhov and Muthen, 2009, p.410-411). Unfortunately, it is easy to see that in bi-factor exploratory models, these conditions are explicitly violated, as all items are expected to load in the general factor. In other words, as no value is expected to be fixed for the general factor, all bi-factor partially specified rotations could result in identification issues [31].

A strategy would be to consider ensuring the identification of the $\mathbf{\Lambda}_{GRP}$ submatrix as to approximate identification of the structure. Following this strategy would allow researchers to get closer to the neighbourhood of identified solutions. An approach similar to the one presented in Myers et al., (2014, footnote 4), where targeted values are changed based on the unrotated solution until identifications are met, could be useful in bi-factor applications. Nevertheless, the aforementioned conditions are only sufficient, but not necessary conditions. However, the impact of non-identifiability issues on the quality of the rotated solutions is still debated in the literature. To date, bi-factor identification issues have only been linked to different characteristics of empirical undetermination. Empirical undetermination implies that parameters cannot be estimated due to the sample data characteristics (Asparouhov and Muthén, 2009; Chen and Zhang, 2018).

[31] ten Berge (2006) identified that a similar problem would occur for the fully-specified target rotation in the simplest structure bi-factor case (p.205). The issue highlighted by the author might be resolved using a strategy similar to the last eigenvalue sign determination applied in Kabsch-Umenaya algorithm applied in shape analysis or other solutions to the more general "Whaba's problem" in satellite attitude determination (Crosilla et al., 2019).

1.7 Bi-factor Models Beyond Factor Loading Recovery

Throughout the previous sections, the evaluation of the different methods for estimating bi-factor models has been largely focused on how to recover the factor loading matrix (i.e., $\mathbf{\Lambda}_{BF}$). Indeed, $\mathbf{\Lambda}_{BF}$ often represents the parameters of most theoretical and practical interest. Moreover, any bias introduced in $\mathbf{\Lambda}_{BF}$ would result in biased parameter estimation in any other model or computation in which they are included (Reise et al., 2018). Due to its prominence, distinct approaches towards understanding how well a factor loading pattern is recovered have been proposed. Most authors favoured using Tucker's factor congruence, as defined by (Lorenzo-seva and ten Berge, 2006). Factor congruence, which represents the cosine of angles between two vectors, provides a standardized similarity measure between two factors, is easy to interpret (its value range are between $\pm$ 1, as in the correlation coefficient) and constitute one of the most widely applied statistics in factor analysis studies. Nevertheless, other authors suggest that could be of benefit considering statistics such as the root mean square, the absolute mean difference, or estimation when assessing factor recovery (Lorenzo-seva and ten Berge, 2006).

Be that as it might be, one of the strengths of the bi-factor model is the extent that it provides additional key information to the researchers. This information is one of the reasons why researchers often prefer the bi-factor model to, for example, the correlated-factors. The most common studied statistics for the exploratory bi-factor model are (Rodriguez et al., 2015, 2016):

1. **Understanding essential unidimensionality**. Researchers are frequently interested in understanding the extent that a general factor accounts for the estimated common variance in the structure, and whether the bi-factor model is set to provide relevant information beyond a unidimensional model. To this end, the preferred statistic in the literature is The Explained Common Variance (i.e., ECV). The ECV is defined as:

$$ECV = \frac{\sum \lambda_G^2}{\sum \lambda_G^2 + (\sum \mathbf{\Lambda}_{GRP})^2} \tag{1.16}$$

   which represents the degree of essential unidimensionality in a bi-factor structure, or the extent that variance explained to the general factor represents the total modelled variance in an orthogonal bi-factor structure (Reise, 2012; Reise et al., 2010; Rodriguez et al., 2015). It should be noticed that the interpretation of ECV values should be conditional on the value of other relevant statistics (Reise et al., 2013). In their systematic review, (Rodriguez et al., 2015) found that, in average, reviewed bi-factor models presented ECV values close to .60, meaning that *"the general factor routinely*

accounted for well over half of the common variance" (Rodriguez et al., 2015, p.231). A complementary statistic to the ECV is the so-called "Parameter Bias", which represent the difference between
$lambda_G$ when estimating a bi-factor or a unidimensional model (Rodriguez et al., 2015).

2. **Assessing general and group factor scores**. An alternative question often encountered in bi-factor modelling is the extent that total score variance is due to the presence of a single general factor. In other words, the extent that reliable model-estimated variance in the total scores is due to the general factor (Reise et al., 2013). This question is often assessed using the omega hierarchical statistic (i.e. ω_H). ω_H is defined as:

$$\omega_H = \frac{(\sum \lambda_G)^2}{(\sum \lambda_G)^2 + (\sum \mathbf{\Lambda}_{GRP})^2 + \sum(1 - \mathbf{\Lambda}_{i\cdot})} \tag{1.17}$$

The squared-root of ω_H represent the correlation between the general factor and observed standardized scores. ω_H values indicate the extent that individual differences in the total scores reflect differences in general or group factor sources of variance. Remarkably, ω_H is related to the omega (i.e., ω coefficient). Omega represents the proportion of variance in total scores that is due to all sources of variances (i.e., general plus group factors). Indeed, ω is defined as:

$$\omega = \frac{(\sum \lambda_G)^2 + (\sum \mathbf{\Lambda}_{GRP})^2}{(\sum \lambda_G)^2 + (\sum \mathbf{\Lambda}_{GRP})^2 + \sum(1 - \mathbf{\Lambda}_{i\cdot})} \tag{1.18}$$

Thus, the difference between ω_H and ω is the inclusion of the variance of total scores that is explained by group factors in the numerator of the equation. ω_H has played a relevant role in bi-factor modelling. ω_H has been identified as a relevant statistic, which has been usually approximated by the SL approach (Zinbarg and Alden, 2015; Zinbarg et al., 2007).

3. **Inspecting latent variables in SEM context**. As bi-factor models start playing a substantive role in structural equation modelling, the consequences of correctly applying bi-factor models have started to be considered (Reise et al., 2018). In this context, researchers might be interested in assessing the properties of the estimated factor scores in the general and group factors. A relevant statistic to this end would be the construct replicability index (Hancock and Mueller, 2001; Rodriguez et al., 2016). This index, which is often called H, represents a ratio between explained and unexplained variance by a given factor. In H, each indicator score is weighted by its factor loading in the

given factor to maximize the reliability of the total score (Aguirre-Urreta et al., 2018). For the general factor, H is defined as:

$$H = (1 + (\sum \frac{\lambda_G^2}{1 - \lambda_G^2})^{-1})^{-1} \tag{1.19}$$

where scores over .80 have been suggested as indicators of a factor being well-represented by a set of items (Ferrando and Lorenzo-Seva, 2017a; Ferrando and Navarro-González, 2018; Rodriguez et al., 2015). H has some interesting properties, seen as superior to traditional reliability estimates from a conceptual standpoint, such as not being affected by the sign of the loadings, always being bounded by the value of the largest squared factor loading (Hancock and Mueller, 2001). Moreover, H values are connected with some measures of factor indeterminacy, which has also been suggested as a useful indicator of factor score estimates' quality (Ferrando and Lorenzo-Seva, 2017a; Rodriguez et al., 2015, 2016).

Remarkably, all these statistics are secondary statistics estimated over the bi-factor pattern matrix. The extent that those loadings are biased (e.g., by approaching a complex structure using an SL transformation), all the reviewed indicators would be biased too. Unfortunately, there is scarce literature regarding how the different algorithms and approaches towards exploratory bi-factor modelling could affect each statistic. Lastly, alternative statistics such as the Percentage of Uncontaminated Correlations (i.e., PUC; Reise, Scheines, Widaman & Haviland, 2013) or omega hierarchical subscale are only appropriately defined for confirmatory bi-factor models, so they are not considered here. In other words, its computation requires researchers to assume a simple bi-factor structure, where the primary loading of each item is requested to be defined, and/or cross-loadings are not taken into account when estimating them. Accordingly, further research is needed to adapt these indicators to the exploratory case.

1.8 Thesis Contributions and Developments

The ultimate goal of this dissertation project was to advance solutions to the methodological challenges posed by the estimation of bi-factor exploratory factor models. Particularly, this thesis will aim to improve available algorithms for performing bi-factor partially specified target rotation while exploring the consequences of these improvements beyond the estimation of the factor loadings matrix. In this sense, this dissertation intended to shed light on the benefits and drawbacks of the different algorithms available, providing creative and novel solutions to overcome some of the limitations identified in the literature. This dissertation is hoped to help to close the gap between psychometricians and practitioners by providing cases of study and software ready to be used by those researchers without specialized knowledge in the topic. Lastly, this dissertation was developed with the hope of encouraging other fellow colleagues to explore the potential impact of research in exploratory methods in factor analysis, highlighting the potential benefits of exploring new uses and applications of factor rotation.

1.8.1 Iterations of Partially Specified Target Matrices: Application to the Bi-factor Case

In Chapter 2, the iterative refinement of the partially specified target rotations will be investigated. A new algorithm, the Iterative Target Rotation based on the Schmid-Leiman solution (i.e., SLi) is then proposed. Additionally, this study was complemented by studying the performance of alternative methods available in the literature, namely bi-quartimin, bi-geomin, and the classical bi-factor partially specified target rotation based on Schmid-Leiman.

1.8.2 Improving Bi-factor Exploratory Modelling: Bi-factor Rotation based on Loading Differences

In Chapter 3, a new strategy for estimating empirical cut-off points for distinguishing near-zero loadings to be minimized during a partial target rotation will be explored and applied to the bi-factor case. A new algorithm, the Empirical Iterative Target Rotation based on Schmid-Leiman solution (i.e., SLiD) is proposed there. This empirical cut-off will be evaluated in realistic conditions where a high number of cross-loadings are expected, showing good statistical properties. Additionally, structures including mixtures of strong and weak group factors are studied in the context of bi-factor models for the first time.

1.8.3 Searching for G: A New Evaluation of SPM-LS Dimensionality

Chapter 4 will present an application of the algorithms presented in Chapters 2 and 3 to the evaluation of the scores derived from a brief intelligence questionnaire. The use of the methods developed in previous chapters allowed to obtain evidence of the presence of additional factors, and the revelation that while this test could be considered to be essentially unidimensional, the presence of a relevant group factor for the last half of items was observed. Accordingly, the theoretical and practical consequences of these findings were discussed in length. Noteworthy, this study was awarded the 2019 Travel Award from the Journal of Intelligence.

1.8.4 On General Factor Reliability: A Comparison of Exploratory Bi-factor Analysis Algorithms

Chapter 5 presents evidence of the consequences of choosing a BEFA algorithm when estimating the omega hierarchical statistic. Furthermore, two new algorithms, the Direct Bi-factor and Direct Schmid-Leiman methods, were additionally considered to study the performance for recovering three types of structures: full-rank bi-factor, second-order (i.e., rank-deficient bi-factor models) and a factor model without a general factor. Additionally, this study provided additional evidence of the functioning of classical methods (bi-geomin, bi-quartimin and SL) under these novel conditions. This research is complemented by studying the functioning of each method in several classical bi-factor empirical datasets.

1.8.5 Bi-factor Structural Equation Modelling Done Right: An application of the SLiD Algorithm

Chapter 7 will explore the consequences of specifying a correct bi-factor measurement model within an ESEM model via a novel estimation of a bi-factor structure for the Generic Conspiracionist Beliefs Scale. In this article a new Shiny application for computing and SLiD-based target matrix to be applied in *Mplus* will be presented, so practitioners could benefit from the application of the modern BEFA algorithms in this context.

Chapter 2

Iteration of Partially Specified Target Matrices: Application to the Bi-Factor Case

This study was published in the journal *Multivariate Behavioral Research*. The following pages include this publication. This article was first published in April 4, 2017. The article can be accessed at https://www.tandfonline.com/doi/abs/10.1080/00273171.2017.1301244

MULTIVARIATE BEHAVIORAL RESEARCH
2017, VOL. 0, NO. 0, 1–14
http://dx.doi.org/10.1080/00273171.2017.1301244

Routledge
Taylor & Francis Group

Iteration of Partially Specified Target Matrices: Application to the Bi-Factor Case

Francisco J. Abad[a], Eduardo Garcia-Garzon[a], Luis E. Garrido[b], and Juan R. Barrada[c]

[a]Autonomous University of Madrid; [b]Universidad Iberoamericana; [c]University of Zaragoza

ABSTRACT
The current study proposes a new bi-factor rotation method, Schmid-Leiman with iterative target rotation (SLi), based on the iteration of partially specified target matrices and an initial target constructed from a Schmid-Leiman (SL) orthogonalization. SLi was expected to ameliorate some of the limitations of the previously presented SL bi-factor rotations, SL and SL with target rotation (SLt), when the factor structure either includes cross-loadings, near-zero loadings, or both. A Monte Carlo simulation was carried out to test the performance of SLi, SL, SLt, and the two analytic bi-factor rotations, bi-quartimin and bi-geomin. The results revealed that SLi accurately recovered the bi-factor structures across the majority of the conditions, and generally outperformed the other rotation methods. SLi provided the biggest improvements over SL and SLt when the bi-factor structures contained cross-loadings and pure indicators of the general factor. Additionally, SLi was superior to bi-quartimin and bi-geomin, which performed inconsistently across the types of factor structures evaluated. No method produced a good recovery of the bi-factor structures when small samples ($N = 200$) were combined with low factor loadings (0.30–0.50) in the specific factors. Thus, it is recommended that larger samples of at least 500 observations be obtained.

KEYWORDS
Bi-factor rotation; exploratory factor analysis; Schmid-Leiman; target rotation

The use of bi-factor analysis has dramatically increased in the last decade (e.g. Chen, West, & Sousa, 2006; Reise, 2012). One of the reasons for this rise in popularity is the ability of these models to separate the latent sources of common variance by their degree of broadness, from the more general to the more specific. Bi-factor models may be used to assess the relative strength and potential usefulness of first-order and higher order factors for multitiered constructs (McDonald, 1999; Zinbarg, Revelle, Yovel, & Li, 2005), as well as to determine the impact of multidimensionality (Reise, Cook, & Moore, 2015). Additionally, they can be used to estimate the relative strength of general and specific factors in the prediction of an external criterion (Bandalos & Kopp, 2013). In its typical form, the bi-factor model has one general factor and a number of specific factors, with the latter explaining common variance that is non-accounted for by the general factor.

When there is insufficient prior knowledge for the domain under investigation, an exploratory approach is needed to uncover possible bi-factor structures (Jennrich & Bentler, 2011). Given the specific restrictions of the bi-factor model, traditional rotation methods (e.g. varimax, oblimin) fail to recover this structure, as they are oriented toward finding simple structures (Reise, Moore, & Maydeu-Olivares, 2011). In order to overcome this challenge, three general strategies have been proposed: (1) exploratory bi-factor analysis using a Schmid-Leiman (SL) orthogonalization (Schmid & Leiman, 1957), which involves a reparameterization of a second-order oblique exploratory factor analysis solution (Yung, Thissen, & McLeod, 1999); (2) SL followed by a target rotation applied to the bi-factor structure (Browne, 2001; Reise et al., 2011); and (3) analytic bi-factor rotation methods such as bi-factor quartimin or bi-factor geomin (Jennrich & Bentler, 2011, 2012).

At the moment there is limited information regarding the performance of the bi-factor rotation methods currently available. On the one hand, many of the previous studies have considered a very specific set of models and conditions (e.g. Asparouhov & Muthén, 2012). On the other hand, the three types of rotation methods have neither been tested under similar conditions, nor directly compared (e.g. Bandalos & Kopp, 2013; Reise et al., 2015), making it difficult to ascertain their relative accuracy and to offer practical guidelines. Furthermore, there is reason to believe that each of these rotation methods has inherent shortcomings in their formulation that may not make them optimal to uncover exploratory bi-factor structures. In light of this, in the current paper we will propose and test a novel strategy that has not been applied to the bi-factor case: the iteration of partially specified target matrices (Moore, Reise, Depaoli, & Haviland, 2015). We call this method SL with iterative target rotation (SLi).

CONTACT Francisco J. Abad fjose.abad@uam.es Facultad de Psicología, Universidad Autónoma de Madrid, 28049 Madrid, Spain.
© 2017 Taylor & Francis Group, LLC

A brief review of the properties and known performance of the aforementioned rotation methods will be presented next, followed by the presentation of the newly proposed bi-factor rotation. In order to better summarize this literature, we will first describe four types of factor structures that may be considered as theoretically and practically relevant for this investigation. Following McDonald (1999, 2000), a factor structure is said to be a perfect *independent cluster* (IC) structure if (1) the specific factors are properly identified (i.e. defined by at least three items on orthogonal structures or by two items on oblique structures) and (2) no cross-loadings are observed. If the former condition is met, but cross-loadings are present on the structure, McDonald (2000, p. 102) named those structures as *independent cluster basis* (ICB) structures. In addition, there can be bi-factor models with variables that represent "pure" indicators of the general factor (i.e. items that have zero loadings on the specific factors) (Mansolf & Reise, 2016), and these will be called *independent cluster pure* (ICP) structures. Finally, bi-factor structures that contain both cross-loadings and pure indicators of the general factor will be referred to as *independent cluster basis pure* (ICBP) structures. The ICBP structures represent realistic factor structures that are often found by practitioners, and are being introduced for the first time in the bi-factor literature in this investigation in order to highlight the relative strengths of the different rotation methods.

Bi-factor rotation methods

Schmid-Leiman rotation (SL)

A brief introduction to the SL transformation (Schmid-Leiman, 1957) is presented in the following section. However, readers interested in a complete description of this procedure and its relationships with the higher order factor models are referred to Yung, Thissen, and McLeod (1999).

The SL method is a multistage procedure. In the first step, the manifest variable correlation matrix ($\mathbf{R}$) is factored with an oblique rotation method (e.g. promax, oblimin, geomin):

$$\mathbf{R} = \mathbf{\Lambda}_0\, \mathbf{\Phi} \mathbf{\Lambda}_0' + \mathbf{\Psi}_0^2 \tag{1}$$

where $\mathbf{\Lambda}_0$ is the loading matrix of the manifest variables on the first-order factors, $\mathbf{\Phi}$ is the first-order factor correlation matrix, and $\mathbf{\Psi}_0^2$ is the diagonal matrix of unique variances for the manifest variables. In a second step, the higher order factor solution is obtained by factoring the lower order factor correlation matrix ($\mathbf{\Phi}$):

$$\mathbf{\Phi} = \boldsymbol{\lambda}_1 \boldsymbol{\lambda}_1' + \mathbf{\Psi}_1^2 \tag{2}$$

where $\boldsymbol{\lambda}_1$ is a vector with the loadings of the first-order factors on the second-order factor, and $\mathbf{\Psi}_1$ is a diagonal matrix with the square root of the unique variances for the first-order factors, which is directly related to $\boldsymbol{\lambda}_1$:

$$\mathbf{\Psi}_1 = [\mathbf{I} - diag(\boldsymbol{\lambda}_1 \boldsymbol{\lambda}_1')]^{1/2} \tag{3}$$

where *diag* indicates that only the diagonal elements from the second-order factor solution are used. Then, the previous model is parameterized as

$$\mathbf{R} = \boldsymbol{\lambda}_g \boldsymbol{\lambda}_g' + \mathbf{\Lambda}_s \mathbf{\Lambda}_s' + \mathbf{\Psi}_0^2 \tag{4}$$

where $\boldsymbol{\lambda}_g\ (= \mathbf{\Lambda}_0 \boldsymbol{\lambda}_1)$ and $\mathbf{\Lambda}_s\ (= \mathbf{\Lambda}_0 \mathbf{\Psi}_1)$ are called SL transformed loadings of the manifest variables on the general and the residualized first-order factors (i.e. after discounting the effects of the general factor), respectively. In this SL parameterization, latent factors are orthogonal and loadings are linearly dependent.

One limitation of the SL method is the assumption that $\boldsymbol{\lambda}_g$ and $\mathbf{\Lambda}_s$ follow only one particular structure. Indeed, all the effects from the general factor to the manifest variables are assumed to be indirect. Because of this, Reise et al. (2015) refer to SL as a "semi-restricted" or *hierarchical* bi-factor model. The more general unrestricted bi-factor model follows the same Equation (4), but $\boldsymbol{\lambda}_g$ and $\mathbf{\Lambda}_s$ do not follow any specific relationship. These structures that do not contain linearly-dependent general and specific factor loadings are known as *non-hierarchical* bi-factor structures.

Reise et al. (2011, 2015) analyzed the performance of SL under IC and ICB population structures. They found that the "semi-restricted" model produced biased estimates of the factor loadings when proportionality constraints were not met in a simple IC structure. In these cases, loadings on the general factor were either underestimated or overestimated depending on the item. In ICB structures, larger distortions were obtained. For example, for items with large cross-loadings (e.g. .40) that broke the proportionality constraints, SL raised an item's communality, causing an overestimation of the loading on the general factor, whereas loadings on the specific factors were underestimated (Reise et al., 2011, 2015).

Schmid-Leiman with target rotation (SLt)

Despite the expected biases of the SL method, Reise, Moore, and Haviland (2010) predicted that the impact of proportionality on real-world data might be negligible when the goal was only to identify *patterns* of salient and non-salient loadings. Following this line of reasoning, Reise et al. (2011) showed that SL was a good method for identifying the pattern of trivial and non-trivial loadings and, thus, SL could be a useful tool for defining a partially specified pattern matrix for a target rotation (Browne, 1972, 2001). Target rotation, also called Procrustean rotation, requires a partially specified target pattern matrix (**B**) in which zeros indicate that the researcher anticipates that the item will not have a salient loading on the factor, while the remaining target values are not specified. The target rotation minimizes the sum of all the squared differences between each specified target value ($b_{ij} = 0$) and the actual corresponding factor loading (λ_{ij}).

In the first step of the Reise et al. (2011) procedure, a cutoff is applied to obtain the target pattern matrix from the SL loading matrix. For example, if the SL loading is greater than or equal to .20, that target pattern loading is marked as an unspecified element, and if the SL loading is less than .20, that target pattern loading is marked as a specified zero. In the second step, the target rotation is applied. Using a cutoff of .15, Reise et al. (2011) showed that for simple IC structures and sample sizes of 500 or above, target misspecification occurred in a low percentage of cases. Additionally, for conditions where the target transformation matrix was correctly specified, the recovery rates were reasonable. However, further research has shown that under ICB structures, cross-loading presence can impair the performance of SL as a tool for correctly specifying a bi-factor target matrix (Reise et al., 2015).

Analytic bi-factor rotations (bi-quartimin and bi-geomin)

Jennrich and Bentler (2011, 2012) developed an analytic rotation method appropriate for reproducing bi-factor structures. They proposed to minimize the following criterion that measures the departure from the bi-factor structure:

$$\mathrm{B}(\mathbf{\Lambda}) = \mathrm{qmin}(\mathbf{\Lambda_2}) = \sum_{i=1}^{I}\sum_{j=2}^{m}\sum_{j'=j+1}^{m} \lambda_{ij}^2\lambda_{ij'}^2 \tag{5}$$

where m is the number of factors, I is the total number of items, and $\mathrm{qmin}\,(\mathbf{\Lambda_2})$ is the bi-quartimin rotation criterion, applied to the pattern matrix after excluding the general factor ($\mathbf{\Lambda_2}$). When a perfect IC structure is obtained, $\mathrm{qmin}\,(\mathbf{\Lambda_2}) = 0$. It follows, therefore, that the bi-quartimin criterion can only be achieved when all cross-loadings in a factor model are zero. It must be noted that $\mathrm{B}(\mathbf{\Lambda})$ does not depend on the first column of $\mathbf{\Lambda}$ (i.e. the general factor), but when $\mathrm{B}(\mathbf{\Lambda})$ is used for rotation, all the columns in $\mathbf{\Lambda}$ are rotated.

Because bi-quartimin rotation attempts to minimize variable complexity by approximating to structures where the items have very low or zero cross-loadings on *all* of the specific factors, it is a method best suited for IC structures. Indeed, Asparouhov and Muthén (2012) simulated IC hierarchical structures under optimal loading size and different number of indicators per specific factor, and concluded that exploratory structural equation models (ESEMs) with bi-quartimin rotation were almost unbiased. However, bi-quartimin is expected to produce biased estimations with ICB structures, and initial studies evaluating its performance in the presence of cross-loadings have supported this expectation (Bandalos & Kopp, 2013).

Another analytic approach developed by Jennrich and Bentler (2012) was the *bi-geomin* rotation method. In this case, the criterion minimized is

$$\mathrm{B}(\mathbf{\Lambda}) = \mathrm{geomin}(\mathbf{\Lambda_2}) = \sum_{i=1}^{I}\prod_{j=2}^{m}\left(\lambda_{ij}^2 + \varepsilon\right)^{1/m} \tag{6}$$

where ε is a small positive value (i.e. .01) needed to make the function differentiable (Browne, 2001; Jennrich & Bentler, 2012).

The bi-geomin rotation method requires only one specific factor loading of zero per item in order to accomplish the criterion (i.e. $\mathrm{B}(\mathbf{\Lambda}) = 0$). Thus, this method attempts to minimize variable complexity by approximating to structures that have one zero-element per row in the pattern matrix of the specific factors. Because of this property, Jennrich and Bentler (2012) expected bi-geomin to have better functioning in the presence of cross-loadings. In this line, Mansolf and Reise (2016) showed the theoretical superiority of bi-geomin to bi-quartimin with ICB structures, an advantage that is borrowed from the superior performance of geomin over quartimin with these structures. Additionally, in a preliminary simulation study Bandalos and Kopp (2013) found that bi-geomin rotation provided a good recovery of ICB structures, whereas bi-quartimin failed to recover the true factor structure in these conditions. Nevertheless, for IC structures, higher samples sizes were necessary (e.g. 2,500) in order for bi-geomin to achieve a correct solution.

There are some additional issues regarding the performance of the analytic bi-factor rotations that should be noted. Firstly, as the general factor is not rotated explicitly, both rotation methods are prone to local minima solutions (Mansolf & Reise, 2016). Indeed, these analytic bi-factor rotations can be conceptualized as a mixture of two factor models: a one-factor model defined by the general factor and an m–1 factor model, where m–1 is the number of specific factors. Depending on the starting values, different variance might be shifted to the general factor, and local minima solutions may be obtained. Also, Mansolf and Reise (2016) warn that the analytic bi-factor rotations will tend to shift as much variance onto the general factor as possible, leading in certain cases to the collapse of the specific factors (i.e. for smaller loadings on the specific factors).

Secondly, analytic bi-factor rotations "break down" when the SL constraints are met (Mansolf & Reise, 2016). That is, when there is a perfect linear dependence between the general and specific factor loadings, a first-order model of m–1 factors can perfectly represent a bi-factor structure of m factors, thus making the latter an overfactored or overparameterized model that can produce Heywood cases and other estimation problems. Therefore, the analytic factor rotations can perform poorly when the SL constraints nearly hold (Mansolf & Reise, 2016).

Schmid-Leiman with iterative target rotation (SLi)

Moore et al. (2015) recently proposed, based on Browne (2001), a method for exploratory factor rotation grounded on the iteration of partially specified factor structures (ITR). In the ITR method, one begins by performing a standard factor rotation and subsequently defines a partially specified, empirically informed target matrix based on the results of this rotation. For this task, a pre-specified loading cutoff criterion (e.g. .20) is established. Then, an iterative search procedure is used to update the target matrix until convergence is reached.

The ITR rotation strategy appears to be particularly useful for data that have a complex structure, such as those with multiple cross-loadings (Moore et al., 2015). This is due to the iterative nature of the method, which has the potential to solve or help ameliorate the problems of using an initially misspecified target matrix. Indeed, Moore et al. (2015) analyzed the performance of ITR with IC and ICB first-order factor structures and found that ITR always outperformed the "one-shot" classical rotations (e.g. quartimin), especially when more cross-loadings were present.

Even though ITR is a promising strategy for factor rotation, it has yet to be applied to the bi-factor case, where a direct generalization of the method can be made. Therefore, we propose in the current paper the use of ITR in conjunction with SL, and call it the SLi method. SLi bi-factor rotation, which can also be considered as an extension of Reise et al. (2011), starts with an initial SL rotation and the obtained rotated matrix is used to specify an initial target matrix (similar to the SLt method). Then, after the target rotation is performed, this new rotated matrix is used to build a new updated target matrix, and the target rotation is again performed. The procedure is repeated in this fashion until the pattern of specified zeros in the target matrix corresponds to the non-salient loadings (e.g. those <.20) in the latest estimated pattern matrix. Due to its iterative approach, SLi should be less affected than SL and SLt by the presence of cross-loadings and pure indicators of the general factor. Moreover, it is expected to be more robust to the issues that particularly affect analytic bi-factor rotations (e.g. local minima, factor collapse, linear dependence of the general and specific loadings), as a result of it being an SL-based method.

Goals of the current study

The present research had two main goals: (1) to evaluate the performance of the newly proposed bi-factor rotation method, SLi, and (2) to compare it with the four bi-factor rotations currently in use, SL, SLt, bi-quartimin, and bi-geomin. As described earlier, the latter four rotation methods had not been studied together, and the current information on their performance was scarce. Further, each of these methods was known or expected to have important shortcomings for particular types of factor structures (see "Bi-factor rotation methods" section), and there was reason to believe that due to its iterative use of targets SLi would provide a better and more consistent performance, in particular for the more complex structures.

In order to achieve the stated goals, a Monte Carlo simulation study was carried out with the manipulation of a large set of variables that were known to affect the performance of the rotation methods. Also, an empirical application of the five bi-factor rotations with a Quality of Life data set (Chen et al., 2006) was undertaken. It should be noted that only bi-factor structures with *orthogonal* factors were considered. As argued by Morin, Arens, and Marsh (2016), these models: (a) ensure interpretable results, and (b) are the most common form of bi-factor methods, with well-known practical applications (e.g. omega reliability coefficient).

Method

The current study considered a comprehensive set of factors and factor levels for the bi-factor models. The following seven variables were manipulated using Monte Carlo methods: (1) sample size (N: 200, 500, 2,000); (2) number of variables per specific factor (VAR.SF: 4, 5, 6); (3) number of specific factors (NUM.SF: 4, 5, 6); (4) presence of cross-loadings on the specific factors (CROSS.SF: no, yes); (5) factor loadings on the specific factors (LOAD.SF: low, medium, high); (6) factor loadings on the general factor (LOAD.GF: low, medium, high); and (7) presence of pure indicators of the general factor (PURE.GF: no, yes). Therefore, the simulation was based on a $3 \times 3 \times 3 \times 2 \times 3 \times 3 \times 2$ factorial design, for a total of 972 conditions.

The factor loadings had ranges from .30 to .50 for the low condition, from .40 to .60 for the medium condition, and from .50 to .70 for the high loading condition. In each case, the loadings for the IC structures were generated with equal increments between loadings under the

Table 1. Examples of the factor loadings and communalities simulated according to the type of structure.

	Independent cluster (IC)						Independent cluster basis (ICB)						Independent cluster pure (ICP)						Independent cluster basis pure (ICBP)					
Item	gf	sf1	sf2	sf3	sf4	h^2	gf	sf1	sf2	sf3	sf4	h^2	gf	sf1	sf2	sf3	sf4	h^2	gf	sf1	sf2	sf3	sf4	h^2
1	.35	.30				.21	.39	.30				.24	.35	.30				.21	.42	.30				.27
2	.34	.37				.25	.47	.37				.36	.54	*.01*				.29	.59	*.01*				.35
3	.39	.43				.34	.41	.43				.35	.31	.43				.29	.30	.43				.28
4	.30	.50				.34	.10	.41	.40			.34	.47	.50				.47	.13	.41	.40			.35
5	.37		.30			.22	.34		.30			.21	.34		.30			.21	.41		.30			.26
6	.45		.37			.33	.37		.37			.27	.61		*.01*			.37	.49		*.01*			.24
7	.41		.43			.35	.35		.43			.31	.38		.43			.33	.47		.43			.41
8	.49		.50			.49	.16		.41	.40		.36	.45		.50			.45	.40		.41	.40		.49
9	.50			.30		.34	.42			.30		.27	.37			.30		.22	.45			.30		.29
10	.31			.37		.23	.50			.37		.38	.59			*.01*		.35	.51			*.01*		.26
11	.42			.43		.36	.31			.43		.29	.33			.43		.29	.38			.43		.33
12	.47			.50		.47	.35			.41	.40	.45	.42			.50		.43	.27			.41	.40	.40
13	.38				.30	.23	.49				.30	.33	.41				.30	.26	.43				.30	.28
14	.43				.37	.32	.38				.37	.28	.62				*.01*	.38	.52				*.01*	.27
15	.33				.43	.29	.46				.43	.40	.30				.43	.28	.34				.43	.30
16	.46				.50	.46	.33	.40			.41	.44	.43				.50	.44	.41	.40			.41	.50
Avg.						.33						.33						.33						.33

Note. gf = general factor; sf = specific factor; h^2 = communality; Avg. = average. IC: no cross-loadings on the specific factors and no pure indicators of the general factor. ICB: cross-loadings but no pure indicators. ICP: pure indicators but no cross-loadings. ICBP: both cross-loadings and pure indicators; Cross-loadings appear underlined; Near-zero loadings in the specific factors appear in italics.

specified range (e.g. the loadings for a factor containing three items in the high range condition were .5, .6, and .7), When cross-loadings were present, the last item for each specific factor had a cross-loading of .40 in the next specific factor. In order to hold constant the communality of the item after adding the cross-loading, a small value was subtracted from each of the remaining non-zero item loadings. For the condition with pure indicators of the general factor, the item in the middle position of each specific factor (e.g. item 2 for a 4-item factor, item 3 for a 5-item factor) had a near-zero loading of .01 in its corresponding specific factor. Here, the loading of the pure item in the general factor was increased so as to maintain the communality equal to what it was before its loading on the specific factor was approximated to zero. An example of population values for the four types of structures simulated (IC, ICB, ICP, ICBP) is presented in Table 1.

Data generation

For each of the simulated conditions, 50 sample data matrices were simulated according to the common factor model procedure. First, the reproduced population correlation matrix (with communalities in the diagonal) was computed

$$\mathbf{R_R} = \mathbf{\Lambda \Phi \Lambda^T} \tag{7}$$

where $\mathbf{R_R}$ is the reproduced population correlation matrix, $\mathbf{\Lambda}$ is the population factor loading matrix, and $\mathbf{\Phi}$ is the population factor correlation matrix.

The population correlation matrix $\mathbf{R_P}$ was then obtained by inserting unities in the diagonal of $\mathbf{R_R}$, thereby raising the matrix to full rank. The next step was performing a Cholesky decomposition of $\mathbf{R_P}$, such that

$$\mathbf{R_P} = \mathbf{U^T U} \tag{8}$$

where $\mathbf{U}$ is an upper triangular matrix. The sample matrix of continuous variables $\mathbf{X}$ was subsequently computed

$$\mathbf{X} = \mathbf{ZU} \tag{9}$$

where $\mathbf{Z}$ is a matrix of random standard normal deviates with rows equal to the sample size and columns equal to the number of variables.

Accuracy criteria

The accuracy of the rotation methods in the recovery of the population structure was evaluated according to Tucker's congruence coefficient (c.c.; Tucker, 1951)

$$\text{c.c.}_{jj} \frac{\sum_{i=1}^{I} \hat{\lambda}_{ij}\lambda_{ij}}{\sqrt{\sum_{i=1}^{I} \hat{\lambda}_{ij}^2 \sum_{i=1}^{I} \lambda_{ij}^2}} \tag{10}$$

where $\hat{\lambda}_{ij}$ is the estimated loading, λ_{ij} is the population loading, I is the total number of items, i is the item number, and j is the factor number.

The congruence coefficient is an index of similarity between factors that has boundaries of −1 and 1. A congruence coefficient in the range of .85–.94 corresponds to a *fair* similarity between factors, while a coefficient of .95 or higher indicates a *good* level of similarity such that the factors can be considered equal (Lorenzo-Seva, & ten Berge, 2006). The procedure used to align the estimated factors with the population factors before computing the congruence coefficient was as follows.

Firstly, in each sample, the direction of an estimated factor was reverted if its average factor loading was negative, as no true population structure presented negative factor loadings. Secondly, all the possible factor order permutations were computed, retaining the solution that minimized the average absolute deviation between the estimated and the true solutions. Thirdly, an estimated factor was reversed in the final solution if its factor congruence coefficient was negative. All factor analyses were performed in the R environment using the unweighted least squares estimator. In order to obtain the correlated factors solution needed for the initial step of the SL methods a geomin rotation was carried out. The analytic bi-factor rotations bi-quartimin and bi-geomin were performed applying the gradient projection algorithm implemented in the *GPArotation package* (Bernaards & Jennrich, 2005). For each sample, the solution selected for these rotation methods was the one that produced the lowest discrepancy function from a total of 10 random starts. In the case of the SLi method, Moore et al. (2015) reported that ITR rotation converged within 7 iterations; for the current study, a maximum of 20 iterations were computed. In addition, loadings lower than .20 were specified as zeros in the target matrices of the SLt and SLi methods. Analyses of variance (ANOVAs) were carried out with the IBM SPSS Statistics v. 20 program. According to Cohen (1988), partial eta squared (η_p^2) effect sizes of .01 represent small effects, .06 medium effects, and .14 or more, large effects.

Results

Monte Carlo simulation

An overall assessment of the accuracy of the bi-factor rotation methods is presented in Table 2, which includes the average congruence coefficients across the levels of the independent variables and in total. Additionally, and in order to better understand the performance of the methods, separate ANOVAs were computed for each method where the congruence coefficient was the dependent variable and the manipulated factors were the between-subjects independent variables. The effect sizes resulting from the ANOVAs are shown in Table 3. To limit the number of results shown, only those interactions that attained a large effect size for at least one of the methods were included in Table 3.

Table 2. Average congruence coefficients for the rotation methods across the manipulated variables.

Variable / Level	SL	SLt	SLi	Bi-quartimin	Bi-geomin
N					
200	.926	.934	.938	.866	.899
500	.948	**.967**	**.975**	.907	**.955**
2000	**.957**	**.982**	**.992**	.926	**.985**
VAR.SF					
4	.924	.944	**.959**	.882	.940
5	.944	**.963**	**.970**	.903	.948
6	**.963**	**.976**	**.976**	.914	**.951**
NUM.SF					
4	.944	**.962**	**.969**	.881	.933
5	.945	**.962**	**.969**	.902	**.951**
6	.942	**.959**	**.967**	.917	**.955**
CROSS.SF					
No	**.966**	**.971**	**.969**	**.958**	.937
Yes	.921	**.951**	**.968**	.842	**.955**
LOAD.SF					
Low	.924	.937	.942	.854	.910
Medium	.948	**.967**	**.976**	.911	**.955**
High	**.959**	**.979**	**.987**	.935	**.973**
LOAD.GF					
Low	.934	**.950**	**.958**	.893	.933
Medium	.944	**.961**	**.968**	.900	.945
High	**.953**	**.972**	**.978**	.908	**.960**
PURE.GF					
No	**.977**	**.974**	**.971**	.922	.927
Yes	.910	.949	**.966**	.878	**.966**
STRUCTURE					
IC	**.984**	**.971**	**.968**	**.953**	.910
ICB	**.969**	**.976**	**.974**	.891	.943
ICP	.948	**.972**	**.971**	**.963**	**.964**
ICBP	.872	.926	**.961**	.794	**.968**
TOTAL	.943	**.961**	**.968**	.900	.946

Note. N = sample size; VAR.SF = variables per specific factor; NUM.SF = number of specific factors; CROSS.SF = cross-loadings in the specific factors; LOAD.SF = loadings in the specific factors; LOAD.GF = loadings in the general factor; PURE.GF = pure indicators of the general factor; SL = Schmid-Leiman; SLt = Schmid-Leiman with target rotation; SLi = Schmid-Leiman with iterative target rotation; IC (independent cluster): no cross-loadings and no pure indicators; ICB (independent cluster basis): cross-loadings but no pure indicators; ICP (independent cluster pure): pure indicators but no cross-loadings; ICBP (independent cluster basis pure): both cross-loadings and pure indicators. Congruence coefficients ≥ .95 appear bolded and underlined.

The results in Table 2 show that SLi was the most accurate and consistent method in recovering the bi-factor structures. The SLi method produced an overall congruence coefficient of .968, which was followed by SLt (c.c. = .961), bi-geomin (c.c. = .946), SL (c.c. = .943), and lastly, bi-quartimin (c.c. = .900). In addition, SLi obtained a congruence coefficient of "good" (≥.95) for 17 of the 19 factor levels that were evaluated (89.5%), thus exhibiting a more consistently accurate performance than the SLt (78.9%), bi-geomin (52.6%), SL (31.6%), and bi-quartimin (5.3%) rotation methods. As expected, the biggest improvements of SLi in comparison to SL and SLt came with structures that contained cross-loadings (c.c.[SLi] = .968 > c.c.[SLt] = .951 > c.c.[SL] = .921) and pure indicators of the general factor (c.c.[SLi] = .966 > c.c.[SLt] = .949 > c.c.[SL] = .910). Indeed, as Table 3 indicates, whereas these variables had a substantial impact in the performance of SLt (η_p^2 [CROSS.SF] = .124; η_p^2 [PURE.GF] = .177), and particularly SL (η_p^2 [CROSS.SF] = .548; η_p^2 [PURE.GF] = .725), their effect was very small for SLi (η_p^2 [CROSS.SF] = .001; η_p^2 [PURE.GF] = .012). In general, all three SL methods were highly affected by the sample size (.283 ≤ η_p^2 ≤ .431) and the loadings in the specific factors (.305 ≤ η_p^2 ≤ .349), but the effect on the

Table 3. Univariate analysis of variance (ANOVA) effect sizes for the rotation methods.

Effect type/variables	SL	SLt	SLi	Bi-quartimin	Bi-geomin
Main effects					
N	**.283**	**.365**	**.431**	**.294**	**.372**
VAR.SF	**.373**	**.192**	.063	.104	.009
NUM.SF	.003	.002	.001	.129	.042
CROSS.SF	**.548**	.124	.001	**.688**	.038
LOAD.SF	**.332**	**.305**	**.349**	**.438**	**.246**
LOAD.GF	.124	.097	.087	.026	.051
PURE.GF	**.725**	**.177**	.012	**.241**	**.148**
Two-way interactions					
VAR.SF × CROSS.SF	**.220**	.102	.009	.052	.001
CROSS.SF × PURE.GF	**.351**	**.189**	.024	**.323**	.024
N × LOAD.SF	**.145**	**.200**	**.257**	.128	**.145**
VAR.SF × PURE.GF	**.296**	.113	.013	.002	.000
Three-way interactions					
VAR.SF × CROSS.SF × PURE.GF	**.168**	.104	.013	.002	.000

Note. N = sample size; VAR.SF = variables per specific factor; NUM.SF = number of specific factors; CROSS.SF = cross-loadings in the specific factors; LOAD.SF = loadings in the specific factors; LOAD.GF = loadings in the general factor; PURE.GF = pure indicators of the general factor; SL = Schmid-Leiman; SLt = Schmid-Leiman with target rotation; SLi = Schmid-Leiman with iterative target rotation. The dependent variable in the ANOVAs was the congruence coefficient. The effect size statistic used was partial eta squared. Large effect sizes (≥ .14) appear bolded and underlined. Only interactions with large effect sizes for at least one method are shown.

number of variables per specific factors was much lower for SLi ($\eta_p^2 = .063$), in comparison to SL ($\eta_p^2 = .373$) and SLt ($\eta_p^2 = .192$).

Regarding the performance of the analytic bi-factor rotations, bi-quartimin produced the worst results of any other method evaluated. This rotation performed poorly for the majority of the factor levels, but was especially sensitive to the cross-loading factor ($\eta_p^2 = .688$), a finding that is line with the theoretical expectations. Bi-geomin, on the other hand, performed much better than bi-quartimin, particularly when the factor structures contained cross-loadings (c.c.[bi-geomin] = .955 >> c.c.[bi-quartimin] = .842) or pure indicators of the general factor (c.c.[bi-geomin] = .966 >> c.c.[bi-quartimin] = .878). In fact, and contrary to the behavior of the other four methods, bi-geomin actually performed better with cross-loadings or pure indicators than without them (Table 2). Despite these results, bi-geomin still produced subpar estimations with small samples of 200 observations (c.c. = .899) or with low loadings in the specific factors (c.c. = .910).

As can be seen in Table 3, the rotation methods were affected by several interactions of the manipulated variables. The two-way interaction of sample size × loadings on the specific factors was the most consistently salient one, producing a large or near-large effect for all of the methods ($.128 \le \eta_p^2 \le .257$). The results of this interaction are plotted in Figure 1 and they show that the performance of all the methods improved with larger samples, but that the improvement was greater for lower loadings in the specific factors. In the case of bi-quartimin and bi-geomin, the accuracy in the recovery of the factor structures was particularly poor when small samples of 200 observations were combined with low factor loadings in the specific factors (c.c. <.85).

An additional interaction that affected particularly the SL ($\eta_p^2 = .168$) and SLt ($\eta_p^2 = .104$) methods was the three-way interaction of variables per specific factor × cross-loadings × pure indicators (Figure 2). The 3 two-way interactions (VAR.SF × CROSS.SF, VAR.SF × PURE.GF, and CROSS.SF × PURE.GF) contained in this three-way interaction all had large or near-large effect sizes for the aforementioned rotations ($.102 \le \eta_p^2 \le .351$), so they were analyzed in the context of the higher order interaction. Also, and in order to better understand the differences in performance between the methods, the three-way interaction was plotted for the other three methods (SLi, bi-quartimin, and bi-geomin), where it had a small or negligible effect ($\eta_p^2 \le .013$). It should be noted that the two-way interaction of cross-loadings x pure indicators did produce a large effect for bi-quartimin ($\eta_p^2 = .323$).

The three-way interaction contained in Figure 2 can be explained as a function of the two-way interaction of cross-loading × pure indicators that in turn interacts with the third factor, number of variables per specific factor.

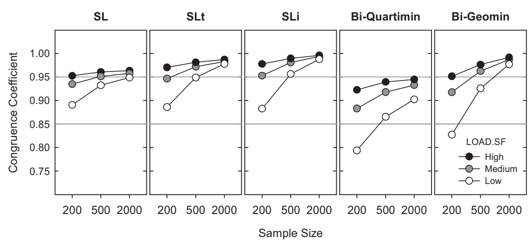

Figure 1. Two-way interaction of N x LOAD.SF with congruence coefficient as dependent variable. *Note*. N, sample size; LOAD.SF, loadings on the specific factors; SL, Schmid-Leiman; SLt, SL target; SLi, SL with iterative target.

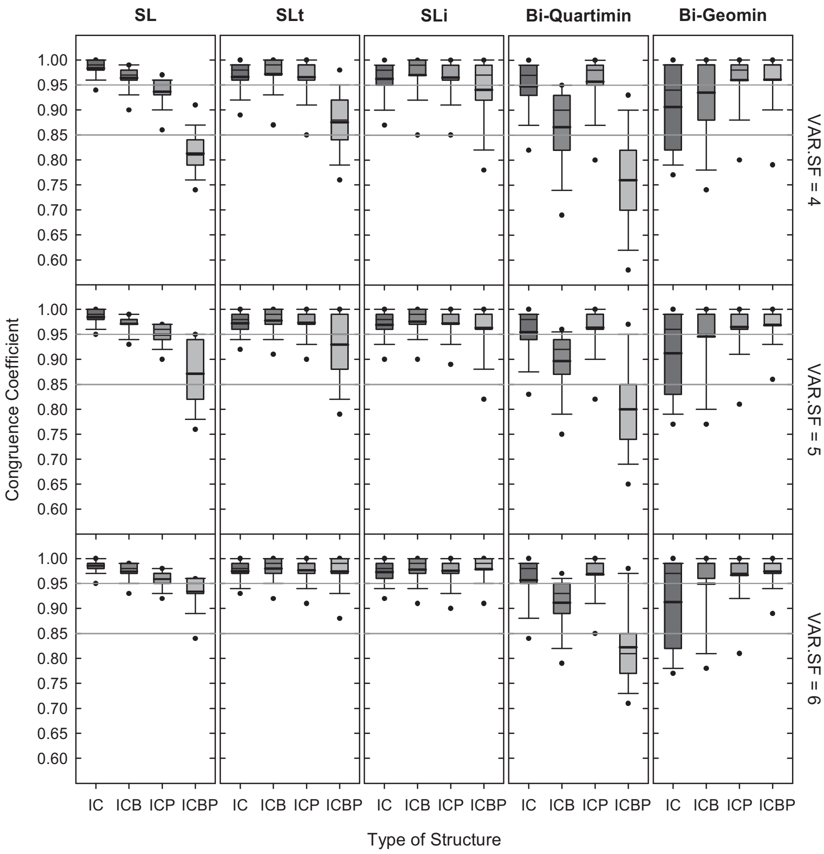

Figure 2. Box plots corresponding to the three-way interaction of VAR.SF x CROSS.SF x PURE.GF with congruence coefficient as dependent variable. *Note*. VAR.SF, variables per specific factor; CROSS.SF, cross-loadings on the specific factors; PURE.GF, pure indicators of the general factor; SL, Schmid-Leiman; SLt, SL target; SLi, SL with iterative target; IC (independent cluster): no cross-loadings and no pure indicators; ICB (independent cluster basis): cross-loadings but no pure indicators; ICP (independent cluster pure): pure indicators but no cross-loadings; ICBP (independent cluster basis pure): both cross-loadings and pure indicators. The thick and thin horizontal lines within each box denote the mean and median. The top and bottom black circles denote the 95th and 5th percentiles.

Firstly, the combined levels of cross-loadings (no, yes) and pure indicators (no, yes) generate the four types of structures considered in this study (IC, ICB, ICP, and ICBP), and their two-way interaction can be clearly seen within each rectangle in Figure 2 for the SL ($\eta_p^2 = .351$), SLt ($\eta_p^2 = .189$), and bi-quartimin ($\eta_p^2 = .323$) rotations. This interaction is evidenced by the substantial differences in accuracy that these rotation methods produce for the four types of structures evaluated. In particular, it can be seen that there was a notably poorer recovery of the ICBP factor structures for these rotations in comparison to their recovery of the IC, ICB, and ICP structures. Additionally, in the case of bi-quartimin, there was also a marked decrease in accuracy for ICB in comparison to the congruence coefficients obtained for IC or ICP. Secondly, the three-way interaction emerged for the SL ($\eta_p^2 = .168$) and SLt ($\eta_p^2 = .104$) methods because the differences in accuracy in the recovery of the four types of structures were greatly diminished as the number of variables per specific factor increased. In contrast, bi-quartimin produced similar results for each level of number of variables per specific factor, which is why the three-way interaction was not salient for this method ($\eta_p^2 = .002$).

The SLi and bi-geomin methods, on the other hand, did not produce important interactions between the factors considered in Figure 2 ($\eta_p^2 \leq .024$ for the two-way interactions and $\eta_p^2 \leq .013$ for the three-way interaction). This was because their recovery accuracy was fairly similar for the four types of structures, regardless of the number of variables per specific factor. The SLi method, in particular, showed the most stable estimations across the IC, ICB, ICP, and ICBP structures, as bi-geomin showed much greater variability in the congruence coefficients that it produced for the IC structures.

Quality of life data set

An empirical study was conducted by factor analyzing a Quality of Life data set (Chen et al., 2006). This Quality of Life data set encompasses 403 observations and 17 items that are hypothesized to reflect a common general factor (Quality of Life) and four specific factors (Cognition, Vitality, Mental Health, and Disease Worry).

There is some controversy regarding the possible bi-factor structure underlying the Quality of Life data set, in particular regarding the third specific factor Mental Health. Using a confirmatory approach, Chen et al. (2006) concluded that this specific factor could be absorbed by the general factor and recommended that it be dropped. In contrast, Jennrich and Bentler (2011) suggested based on a bi-quartimin rotation that it might be retained as two of its items produced salient loadings on this specific factor. For the current study, the five bi-factor rotation methods under investigation were applied to the Quality of Life data set by factorizing the covariance matrix provided in Chen et al. (2006) with the package *psych* (Revelle, 2016). In the case of the SLt and SLi methods, a cutoff of .20 (Jennrich & Bentler, 2011; Moore et al., 2015) was used to distinguish between salient and non-salient loadings for the specification of the target matrices.

The factor loading matrices corresponding to the five bi-factor rotation methods are shown in Table 4. As evidenced by Table 4, the factor loadings for the general factor (Quality of Life) and the first (Cognition) and fourth (Disease Worry) specific factors are consistently high, with no cross-loadings (≥.20) for these items in any of the rotations. Indeed, when congruence coefficients were computed between each pair of rotation methods for these factors, they were extremely high: between .997 and 1.000 for Quality of Life, .983 and 1.000 for Cognition, and between .972 and 1.000 for Disease Worry. In the case of the second specific factor (Vitality), the congruence coefficients between the SL, SLt, SLi, and bi-geomin methods were also especially high (.978 ≤ c.c. ≤ .997); however, they were somewhat lower for the four pairs that contained the bi-quartimin rotation (.936 ≤ c.c. ≤ .957). Interestingly, this factor had the only item ("Feel full of pep?") that achieved a cross-loading of at least .20 for any of the rotations, and the previous simulation results had shown that bi-quartimin was highly affected by the presence of cross-loadings.

As with previous factor analyses of this data set, the greatest differences in factor loadings were obtained for the third specific factor (Mental Health). Here, the congruence coefficients between seven pairs of methods were notably low (.290 ≤ c.c. ≤ .685). The only three congruence coefficients that showed good agreement were between SLt and SLi (c.c. = .958), SLt and bi-geomin (c.c. = .967), and between SLi and bi-geomin (c.c. = .995). It is noteworthy in this case that the factor loadings suggested by SLi and bi-geomin for this specific factor include three items ("Feel downhearted and blue?," "Feel very nervous?," and "Feel so down in the dumps nothing could cheer you up?") that are essentially pure indicators of the general factor, as they produced negligible loadings on the specific factor. If indeed the population structure had these characteristics, the findings would be in line with the simulation results of this study, which showed that SLi and bi-geomin were the two most likely methods to produce accurate recoveries of the factor loadings when pure indicators were present. If researchers are interested in reproducing the presented analysis, the R code necessary for computing the SLi rotation of the Quality of Life data set can be found in the Supplementary Materials of this article.

Discussion

For hierarchically structured constructs that operate at various levels of generality, bi-factor analysis has become an essential modeling technique as a result of its capability to separate the general and specific variances underlying the observed data (Brunner, Nagy, & Wilhelm, 2012). Several rotation methods have been proposed for exploratory bi-factor analysis, including the SL orthogonalization (SL; Schmid & Leiman, 1957), SLt(SLt; Reise et al., 2011), and two analytic bi-factor rotations, bi-quartimin and bi-geomin (Jennrich & Bentler, 2011, 2012). However, at the moment, there is limited information regarding the performance of these rotations under varying data characteristics and in comparison to each other. Furthermore, there are concerns regarding the efficacy of these rotations for certain types of factor structures that are based on their theoretical formulations and the empirical evidence that is available (Bandalos & Kopp, 2013; Mansolf & Reise, 2016; Reise et al., 2011, 2015). Taking into consideration the issues outlined previously, a new bi-factor rotation was proposed in the current study based the use of iterative targets in conjunction with an initial SL orthogonalization: The SLi method. To test the accuracy of this new rotation, an extensive simulation study was undertaken where seven relevant variables were manipulated, thus permitting an in-depth comparison of the accuracy of SLi against the other four bi-factor rotations. The most important findings from this Monte Carlo study will be discussed next, as well as the results obtained with a Quality of Life data set.

Table 4. Bi-factor rotation methods applied to the Chen et al. (2006) quality of life data.

	Schmid-Leiman (SL)					SL with target rotation (SLt)					SL with iterative target rotation (SLi)					Bi-quartimin					Bi-Geomin				
Item	gf	sf1	sf2	sf3	sf4	gf	sf1	sf2	sf3	sf4	gf	sf1	sf2	sf3	sf4	gf	sf1	sf2	sf3	sf4	gf	sf1	sf2	sf3	sf4
diffreas	**.53**	**.67**	.00	.01	− .01	**.55**	**.65**	.00	− .12	− .02	**.59**	**.62**	− .03	− .11	− .04	**.56**	**.64**	.02	.11	.00	**.58**	**.62**	− .02	− .12	− .04
sloract	**.45**	**.47**	.06	− .01	.04	**.45**	**.47**	.07	− .03	.04	**.47**	**.46**	.06	− .03	.03	**.48**	**.45**	.06	.00	.03	**.47**	**.46**	.06	− .03	.03
confsed	**.54**	**.64**	.01	.01	.02	**.54**	**.66**	.03	.08	.04	**.55**	**.64**	.03	.07	.04	**.60**	**.60**	− .02	− .09	.00	**.55**	**.65**	.03	.07	.04
forget	**.44**	**.67**	.00	− .02	− .03	**.44**	**.68**	.02	.03	− .02	**.46**	**.66**	.01	.02	− .02	**.49**	**.64**	− .01	− .06	− .04	**.45**	**.67**	.02	.01	− .02
diffconc	**.56**	**.63**	.00	.04	.00	**.57**	**.61**	.01	− .03	.00	**.60**	**.59**	− .01	− .02	− .01	**.59**	**.58**	.00	.04	.01	**.59**	**.59**	− .01	− .03	− .01
tired	**.65**	.02	**.54**	.04	− .03	**.65**	.05	**.54**	− .02	− .03	**.65**	.03	**.54**	− .01	− .04	**.70**	− .01	**.47**	.00	− .07	**.67**	.02	**.52**	.02	− .06
enrgtic	**.55**	− .01	**.38**	.04	.07	**.55**	.00	**.38**	− .01	.06	**.55**	− .01	**.38**	.00	.06	**.58**	− .04	**.33**	.00	.03	**.57**	− .02	**.36**	.01	.04
wornout	**.65**	.06	**.57**	− .04	.06	**.64**	.09	**.58**	− .09	.06	**.65**	.07	**.56**	− .10	.04	**.68**	.04	**.54**	.00	.02	**.67**	.06	**.55**	− .07	.02
peppy	**.67**	− .03	**.39**	.15	− .02	**.67**	− .02	**.42**	.17	− .01	**.66**	− .02	**.44**	**.20**	.01	**.74**	− .11	**.28**	− .10	− .08	**.67**	− .03	**.40**	**.21**	− .01
atpeace	**.72**	.02	− .04	**.36**	.02	**.77**	− .04	− .06	**.34**	.00	**.73**	− .03	− .02	**.42**	.04	**.80**	− .12	− **.22**	− .04	− .04	**.73**	− .02	− .07	**.41**	.02
feelblue	**.75**	.05	.00	**.31**	.08	**.83**	− .05	− .08	− .04	.01	**.83**	− .08	− .08	.05	.00	**.78**	− .05	− .08	**.29**	.07	**.83**	− .07	− .11	.03	− .02
happy	**.66**	− .03	.06	**.33**	− .02	**.70**	− .07	.05	**.28**	− .03	**.67**	− .07	.09	**.35**	.00	**.74**	− .15	− .10	− .04	− .08	**.68**	− .06	.04	**.35**	− .02
nervous	**.64**	**.24**	.01	.17	.09	**.68**	.19	− .02	− .03	.06	**.69**	.17	− .02	.02	.05	**.66**	.18	− .03	.16	.09	**.69**	.17	− .04	.01	.04
down	**.70**	.13	.05	**.27**	.01	**.81**	.01	− .05	− **.20**	− .10	**.83**	− .03	− .06	− .10	− .12	**.73**	.05	.01	**.42**	.00	**.82**	− .02	− .09	− .12	− .14
afraid	**.66**	.03	− .02	.00	**.60**	**.66**	.03	− .05	− .07	**.60**	**.67**	.00	− .06	− .06	**.59**	**.63**	.02	− .03	.06	**.63**	**.68**	.00	− .07	− .08	**.57**
frust	**.68**	.00	.13	.02	**.44**	**.67**	.02	.12	.02	**.45**	**.67**	.00	.12	.02	**.45**	**.68**	− .02	.09	− .01	**.44**	**.68**	− .01	.10	.02	**.43**
hlthwry	**.61**	− .02	.02	.01	**.52**	**.59**	.01	.02	.08	**.56**	**.59**	.00	.02	.07	**.56**	**.61**	− .04	− .03	− .08	**.54**	**.60**	− .01	.01	.07	**.55**

Note. gf = general factor (Quality of life); sf = specific factor (sf1: Cognition; sf2: Vitality; sf3: Mental health; sf4: Disease worry); diffreas: "Have difficulty reasoning and solving problems?"; sloract: "React slowly to things that were said or done?"; confsed: "Become confused and start several actions at a time?"; forget: "Forget where you put things or appointments?"; diffconc: "Have difficulty concentrating?"; tired: "Feel tired?"; enrgtic: "Have enough energy to do the things you want?"; wornout: "Feel worn out?"; peppy: "Feel full of pep?"; atpeace: "Feel calm and peaceful¿'; feeleblue: "Feel downhearted and blue?"; happy: "Feel very happy?"; nervous: "Feel very nervous?"; down: "Feel so down in the dumps nothing could cheer you up?"; afraid: "Were you afraid because of your health?"; frust: "Were you frustrated about your health?"; healthwry: "Was your health a worry in your life?"; The dotted lines indicate the item groupings according to the theoretical dimensions. Factor loadings ≥ .20 in absolute value appear bolded and underlined.

Main findings

The results pertaining to the SL rotation showed that it produces the highest levels of accuracy of any method for IC structures, but that it is much less effective with complex structures, in particular those that combine cross-loadings with pure indicators of the general factor. These results are in line with the theoretical expectations, as the latter structure presents the greatest departure from the hierarchical model that is the basis for its formulation. In particular, pure indicators constitute severe violations of the hierarchical model assumed by SL rotation where the general and specific factor loadings are linearly dependent or proportional. Previous research had also shown that SL produces biased estimates of the factor loadings in the presence of cross-loadings (Reise et al., 2011, 2015). Additionally, the number of variables per specific factor, the factor loadings in the specific factors, and the sample size, affect the performance of SL, which produces substantially lower levels of accuracy when these variables have smaller values.

Using a one-shot target rotation with the SLt method improves the performance of SL slightly for structures with cross-loadings and more substantially for structures with pure indicators. In particular, SLt is advantageous for very complex structures that combine cross-loadings with pure indicators. It appears, therefore, that SL can be a useful tool to define a partially specified pattern matrix for target rotation, as suggested by Reise et al. (2010, 2011), and that using the SLt method can correct some of the misspecifications that SL produces with these complex structures. However, the performance of SLt with structures that combine cross-loadings and pure indicators is still very variable and mostly below the levels that are considered to represent a good factorial recovery. This is true, especially in those cases where there are also a small number of variables per specific factor. Regarding the other independent variables, the accuracy of SLt is affected by the sample size and the loadings in the specific factors in a similar way as the SL method.

The SLi method was introduced in this study with the aim of improving the performance provided by SLt with the more complex structures, and the findings from the Monte Carlo simulation suggest that indeed it is capable of accomplishing this goal. The performance of SLi is nearly identical to that of SLt for the majority of the data structures, except for the most complex ones that combine cross-loadings and pure indicators. In these cases, SLi is much less variable and produces substantially higher levels of accuracy than SLt, in particular when these types of structures also have a small number of variables per factor. Therefore, it can be concluded that using iterative targets is a useful strategy for bi-factor rotation, particularly when the population structures have diverse departures from the IC model. These findings extend those of Moore et al. (2015), which had proposed and evaluated the use of iterative targets with first-order factor models. In general, the evidence suggests that SLi is the most consistent and accurate of the bi-factor rotations considered here.

The performances of the two analytic bi-factor rotations, bi-quartimin and bi-geomin, are distinctly different. Bi-quartimin produces good accuracy levels for structures that contain ICs or that deviate from them only due to pure indicators of the general factor. However, its performance is notably poor when cross-loadings are introduced and even worse when they are combined with pure indicators. The poor results of bi-quartimin with cross-loadings are in line with Bandalos and Kopp (2013) and reflect the major theoretical shortcoming of the quartimin rotation: it attempts to minimize variable complexity by searching for structures where the items have very low or zero cross-loadings on *all* of the factors. Here, the evidence suggests that when the population structure deviates from the IC model (as it often does in practice) the impact on the accuracy of bi-quartimin is extreme. Bi-geomin, on the other hand, produces a unique pattern of results that is unlike that of the other methods. With IC structures bi-geomin obtains its worst accuracy levels, which are also substantially below the ones of the other methods evaluated. This result was expected, as the bi-geomin criterion is minimized for structures that contain at least *one* non-zero cross-loading. When cross-loadings are introduced, the performance of bi-geomin shows a notable improvement that includes much higher levels of accuracy than bi-quartimin, in line with Mansolf and Reise (2016), but that are still considerably lower than those of the SL-based methods. Surprisingly, bi-geomin produces its best performance with pure indicators, achieving its highest accuracy and that of any method for the most complex structures that combine cross-loadings and pure indicators. This is a unique finding of this study that points to the usefulness of this rotation for these types of structures. Nonetheless, it should be mentioned that bi-geomin is a method that performs considerably different across sample sizes, and that is not really suited for small samples.

From the results of this simulation study it is not possible to determine if the poor performance of the analytic rotations for certain factor structures can be attributed in part to the more correct rotation being contained in a local minimum rather than the global minimum (chosen here). Analytic bi-factor rotations are prone to local minima solutions because the general factor is not rotated explicitly (Mansolf & Reise, 2016). That is, analytic bi-factor rotations utilize a two-stage process where in the first "gradient descent" step only the specific factors are rotated, excluding the general factor. Then, in the second "projection" step the obtained solution is projected in order to obtain a proper factor loading matrix. Thus, the general factor is only rotated implicitly (i.e. by projection to a proper solution). For practical use, it has been suggested that researchers examine the different local minima and global minimum solutions (as there is no mathematical reason to prefer one over the other) and to select the most interpretable one (Asparouhov & Muthén, 2009). However, it is unknown if this process would lead to actually choosing the more correct or replicable solution in practice more often or not.

An empirical study based on a Quality of Life data set (Chen et al., 2006) included in the study appears to support the results of the Monte Carlo simulation. The findings related to the congruence of the factor solutions obtained by the different rotation methods show that for complex factors (those that appear to have items with cross-loadings or pure indicators) bi-quartimin and SL are the methods that have the least agreement with the others, suggesting that their solutions may not be accurate estimations of the population structure. Additionally, for these factors the methods with the highest agreement are SLi and bi-geomin, which in the simulation were the ones that performed the best for structures that contained both cross-loadings and pure indicators. For strong factors that contained items without notable cross-loadings and that had substantial loadings in both the general factor and their respective specific factors, all the methods showed very high agreement with each other.

Limitations

As with any Monte Carlo simulations, the findings of this study are only generalizable to the conditions that were analyzed. Some additional limitations, as well as recommendations for future research, will be addressed next. First, the factor analyses in this study were carried out using the unweighted least squares estimator over Pearson correlations obtained from continuous variables. More research is needed to understand how the bi-factor rotations perform with other estimators (e.g. maximum likelihood, weighted least squares), types of variables (e.g. ordered-categorical), and measures of association (e.g. polychoric correlations). Second, all the simulated structures were balanced, with equal numbers per factor of variables, cross-loadings, and pure indicators. The study of unbalanced structures could provide further insight regarding the accuracy of the bi-factor rotations. Third, iterative targets were evaluated in conjunction with an initial SL orthogonalization based on a geomin rotation for the first-order factor analysis. Additional research is needed to determine how bi-factor rotations with iterative targets would perform with an initial target based on other methods, such as a bi-geomin rotation, or with a SL orthogonalization based on other oblique rotations like oblimin, which is implemented in the SCHMID routine of the *psych* package (Revelle, 2016).

Another issue of importance related to iterative target rotation is the selection of the factor loading cutoff value needed to specify the target matrices. In the present study, a theoretical cutoff value of .20 was used to determine if a loading was to be considered as salient or non-salient. At this moment it is unknown how using other cutoff values would affect the recovery of the bi-factor structures. A possible alternative to this issue could be the use of empirically derived cutoff values, like it is done with promin rotation (Lorenzo-Seva, 1999) or with the standard error method (Moore, 2013). The combination of empirical specifications of the target transformation with empirical cutoff values could ultimately lead to an application of target rotation methods that does not require any additional input from the researcher.

Practical implications

The findings from this study suggest that there are important differences in the levels of accuracy with which the different rotation methods currently available can recover exploratory bi-factor structures. In light of this, it is important for applied researchers to be cognizant of the methods that can best aid them in uncovering these structures, those that should be avoided, and the conditions where none are likely to perform well.

The SLi method, proposed for the first time in this study, was the most accurate and consistent bi-factor rotation across the wide range of conditions that were explored, thus making it one of the methods that can be recommended for applied research. In contrast, the original SL rotation is not recommended due to its notable poorer performance with structures that deviate from the IC model. In the case of the SLt method, its performance was mainly as good as or worse than SL with iterative targets, making the latter the obvious choice for general applied use. Regarding the analytic bi-factor rotations, bi-quartimin is clearly a method that should be avoided due it is markedly poor accuracy across the majority of factor structures. Bi-geomin, on the other hand, can be recommended for cases where the researchers expect complex structures that contain both cross-loadings and pure indicators of the general factor. It should be used with caution, however, because bi-geomin requires larger samples and tends to perform very poorly when the structures contain ICs. Finally, no method is likely to produce a good recovery of the bi-factor structures when small samples (N =

200) are combined with low factor loadings (0.30–0.50) in the specific factors. If this situation is expected, it is recommended that larger samples be obtained in order to offset the detrimental effects of the low item loadings.

Article information

Conflict of interest disclosures: Each author signed a form for disclosure of potential conflicts of interest. No authors reported any financial or other conflicts of interest in relation to the work described.

Ethical principles: The authors affirm having followed professional ethical guidelines in preparing this work. These guidelines include obtaining informed consent from human participants, maintaining ethical treatment and respect for the rights of human or animal participants, and ensuring the privacy of participants and their data, such as ensuring that individual participants cannot be identified in reported results or from publicly available original or archival data.

Funding: This work was supported by Grant PSI2013-44300-P from the Ministerio de Economia y Competitividad (Spain) and the European Social Fund and by Grant FPU15/03246 from the Ministerio de Educación, Cultura y Deporte (Spain).

Role of the funders/sponsors: None of the funders or sponsors of this research had any role in the design and conduct of the study; collection, management, analysis, and interpretation of data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Acknowledgments: The ideas and opinions expressed herein are those of the authors alone, and endorsement by the authors' institutions is not intended and should not be inferred.

References

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, *16*, 397–438.

Asparouhov, T., & Muthén, B. (2012). Comparison of computational methods for high dimensional item factor analysis. *Mplus Technical Report*. Retrieved from http://statmodel.com/download/HighDimension.pdf

Bandalos, D. L., & Kopp, J. P. (2013). *Using the bi-factor model in scale construction*. Paper presented at the annual meeting of the Association for Psychological Science. Washington, DC. Retrieved from http://www.jmu.edu/assessment/research/students/APS_2013_Bandalos.pdf

Bernaards, C. A. & Jennrich, R. I. (2005). Gradient projection algorithms and software for arbitrary rotation criteria in factor analysis. *Educational and Psychological Measurement*, *65*, 676–696. http://dx.doi.org/10.1177/0013164404272507

Browne, M. W. (1972). Orthogonal rotation to a partially specified target. *British Journal of Mathematical and Statistical Psychology*, *25*, 115–120. http://dx.doi.org/10.1111/j.2044-8317.1972.tb00482.x

Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research*, *35*, 111–150. http://dx.doi.org/10.1207/S15327906MBR3601_05

Brunner, M., Nagy, G., & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs. *Journal of Personality*, *80*(4), 796–846. http://dx.doi.org/10.1111/j.1467-6494.2011.00749.x

Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bi-factor and second order models of quality-of-life. *Multivariate Behavioral Research*, *41*(2), 189–225. http://dx.doi.org/10.1207/s15327906mbr4102_5

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York, NY: Routledge Academic.

Jennrich, R.I., & Bentler, P.M. (2011). Exploratory bi-factor analysis. *Psychometrika*, *76*, 537–549. http://dx.doi.org/10.1007/bf02294706

Jennrich, R. I., & Bentler, P. M. (2012). Exploratory bi-factor analysis: The oblique case. *Psychometrika*, *77*, 442–454. http://dx.doi.org/10.1007/s11336-012-9269-1

Lorenzo-Seva, U. (1999). Promin: A method for oblique factor rotation. *Multivariate Behavioral Research*, *34*(3), 347–365.

Lorenzo-Seva, U., & Ten Berge, J. M. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology*, *2*, 57–64. http://dx.doi.org/10.1027/1614-1881.2.2.57

Mansolf, M., & Reise, S. P. (2016). Exploratory bifactor analysis: The Schmid-Leiman orthogonalization and Jennrich-Bentler analytic rotations. *Multivariate Behavioral Research*, *51*(5), 698–717. http://dx.doi.org/10.1080/00273171.2016.1215898

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, N.J.: L. Erlbaum Associates.

McDonald, R. P. (2000). A basis for multidimensional item response theory. *Applied Psychological Measurement*, *24*, 99–114. http://dx.doi.org/10.1177/01466210022031552

Moore, T. M. (2013). *Iteration of target matrices in exploratory factor analysis* (Unpublished doctoral dissertation). University of California, Los Angeles.

Moore, T. M., Reise, S. P., Depaoli, S., & Haviland, M. G. (2015). Iteration of partially specified target matrices: Applications in exploratory and bayesian confirmatory factor analysis. *Multivariate Behavioral Research*, *50*, 149–161. http://dx.doi.org/10.1080/00273171.2014.973990

Morin, A. J., Arens, A. K., & Marsh, H. W. (2016). A bi-factor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling*, *23*, 116–139. http://dx.doi.org/10.1080/10705511.2014.961800

Reise, S. P. (2012). The rediscovery of bi-factor measurement models. *Multivariate Behavioral Research*, *47*, 667–696. http://dx.doi.org/10.1080/00273171.2012.715555

Reise, S. P., Cook, K. F., & Moore, T. M. (2015). Evaluating the impact of multidimensionality on unidimensional item response theory model parameters. In S.P. Reise & D.A. Revicki (Eds.), *Handbook of item response theory modeling applications to typical performance assessment* (pp. 13–40). New York: Routledge.

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bi-factor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment*, *92*, 544–559. http://dx.doi.org/10.1080/00223891.2010.496477

Reise, S. P., Moore, T., & Maydeu-Olivares, A. (2011). Targeted bi-factor rotations and assessing the impact of model violations on the parameters of unidimensional and bifactor models. *Educational and Psychological Measurement*, *71*, 684–711. http://dx.doi.org/10.1177/0013164410378690

Revelle, W. (2016). *psych: Procedures for personality and psychological research*. Evanston, IL: Northwestern University. Retrieved from http://cran.r-project.org/package=psych

Schmid, S., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika*, *22*, 53–61. http://dx.doi.org/10.1007/bf02289209

Tucker, L. R. (1951). *A method for synthesis of factor analysis studies* (Personnel Research Section Report No. 984). Washington, DC: Department of the Army.

Yung, Y., Thissen, D., & McLeod, L. M. (1999). On the relationship between the high order factor model and the hierarchical factor model. *Psychometrika*, *64*, 113–128. http://dx.doi.org/10.1007/bf02294531

Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α, Revelle's β, and McDonald's ωH): Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, *70*, 123–133. http://dx.doi.org/10.1007/s11336-003-0974-7

Chapter 3

Improving Bi-factor Exploratory Modeling

This study was published in the journal *Methodology: The European Journal of Research Methods for the Behavioral and Social Sciences*. The following pages include this publication. This article was first published in May 27, 2019. The article can be accessed at https://doi.org/10.1027/1614-2241/a000163

${protocol}://econtent.hogrefe.com/doi/pdf/10.1027/1614-2241/a000163 - Eduardo Garcia-Garzon <eduardo.garciag@uam.es> - Saturday, July 27, 2019 12:14:13 AM - IP Address:83.34.101.73

Original Article

Improving Bi-Factor Exploratory Modeling

Empirical Target Rotation Based on Loading Differences

Eduardo Garcia-Garzon[1], Francisco José Abad[2], and Luis Eduardo Garrido[3,4]

[1]Departamento de Psicología Social y Metodología, Facultad de Psicología, Universidad Autónoma de Madrid, Spain
[2]Facultad de Psicología, Universidad Autónoma de Madrid, Charlottesville, Spain
[3]Escuela de Psicología, Pontificia Universidad Católica Madre y Maestra, Santo Domingo, Dominican Republic
[4]Department of Psychology, University of Virginia, Charlottesville, VA, USA

Abstract: Bi-factor exploratory modeling has recently emerged as a promising approach to multidimensional psychological measurement. However, state-of-the-art methods relying on target rotation require researchers to select an arbitrary cut-off for defining the target matrix. Unfortunately, the consequences of such choice on factor recovery remain uninvestigated under realistic conditions (e.g., factors differing in their average loadings). Built upon the iterative target rotation based on Schmid-Leiman algorithm (SLi), a novel method is here introduced (SLiD). SLiD settles an empirical, factor-specific cut-off based on the first prominent one-lagged difference of sorted squared normalized factor loadings. SLiD and SLi with arbitrary cut-off (ranging from .05 to .20) performance were evaluated via Monte Carlo simulation manipulating sample size, number of specific factors, number of indicators, and cross-loading magnitude. Results indicate that SLiD performed the best for all conditions. For SLi, and due to the presence of minor factors, smaller cut-offs (i.e., .05) outperformed higher ones (i.e., .20).

Keywords: bi-factor, target rotation, Schmid-Leiman, exploratory factor analysis

Bi-factor models (Holzinger & Swineford, 1937) are regarded today as crucial tools for evaluating psychological constructs, as they allow for the inspection of the simultaneous direct influence of a general and specific factors over a set of indicators (Reise, 2012). Although traditionally studied by means of confirmatory models that conform to a simplest solution for the specific factors, the suitability of such proposals has been recently questioned (Morin, Arens, & Marsh, 2015). This has led to a growing interest on more realistic exploratory bi-factor models, where each indicator is expected to be potentially explained by all factors (Abad, Garcia-Garzon, Garrido, & Barrada, 2017; Mansolf & Reise, 2016; Waller, 2017).

Automated bi-factor target methods have emerged as a compelling alternative for performing exploratory bi-factor analysis (Abad et al., 2017). However, these procedures require researchers to choose a single cut-off to distinguish which factor loadings are expected to be negligible in the final rotated solution (Browne, 1972). Unfortunately, the impact of this decision has not been sufficiently examined (Myers, Ahn, & Jin, 2013; Myers, Jin, Ahn, Celimli, & Zopluoglu, 2015). Furthermore, as the application of a single cut-off could be inappropriate, some authors have suggested the use of factor-specific cut-offs for achieving simple solutions (Lorenzo-Seva, 1999; Trendafilov, 1994).

The current study aims to propose a new method for target specification based on *factor loading differences* to empirically determine factor-specific cut-offs. This new method is applied to the Schmid-Leiman with iterative target rotation (SLi; Abad et al., 2017). SLi is a bi-factor rotation algorithm that has been shown to perform more accurately than both, analytical bi-factor rotations and non-iterative Schmid-Leiman target rotation (Abad et al., 2017). The performance of the proposed method, hereafter SLiD, is therefore tested against that of SLi with a range of predetermined fixed cut-offs (see Appendix).

The rest of the article is structured as follows. First, the Schmid-Leiman orthogonalization and the automatic specification of target matrices are briefly reviewed. Next, an overview regarding the iterative target rotation procedure is presented. In the following section, an introduction to empirical bi-factor target rotation based on loading differences method is depicted. Subsequently, the results from an extensive Monte Carlo simulation comparing the accuracy of SLiD to that of SLi with arbitrary cut-offs are examined. Finally, implications for future research are discussed.

© 2019 Hogrefe Publishing

https://doi.org/10.1027/1614-2241/a000163

${protocol}://econtent.hogrefe.com/doi/pdf/10.1027/1614-2241/a000163 - Eduardo Garcia-Garzon <eduardo.garciag@uam.es> - Saturday, July 27, 2019 12:14:13 AM - IP Address:83.34.101.73

In the remainder of the article, a factor pattern matrix is said to follow a bi-factor model if and only if a general factor (i.e., which directly influences most indicators) along with several specific factors (i.e., which influence distinct groups of indicators within the same set) are present, where the latter explains variance residual to the former, and all factors are orthogonal (Holzinger & Swineford, 1937; Reise, 2012). Noteworthy, an item is always expected to load in the general and in an unrestricted number of specific factors. Therefore, cross-loadings arising from an item reflecting variance from several group factors are to be expected (Morin et al., 2015), a situation that constitutes a deviation from traditional, simple bi-factor structures.

The Schmid-Leiman Transformation

Early proposals for conducting exploratory bi-factor analysis relied on the Schmid-Leiman orthogonalization (SL; Schmid & Leiman, 1957), which transforms a hierarchical solution into a model with the appearance of a bi-factor model. However, the latter model presents linear dependencies between the general and specific factors loadings (Mansolf & Reise, 2016). Indeed, Waller (2017) demonstrates that for any structure defined by three (or more) common factors, an SL solution is not unique (p. 1), and it represents a low-rank solution when compared with a bi-factor structure without such linear constraints.[1] Consequently, a SL solution would only provide unbiased estimates when recovering a true structure presenting such constraints (viz., a hierarchical model; Abad et al., 2017; Reise, Moore, & Maydeu-Olivares, 2011). Unfortunately, SL solutions are often considered as full-rank bi-factor structures and misguidedly applied to compute statistics such as scale reliability. In Waller's (2017) words:

> "the [SL] model sometimes fools researchers into believing that the SL transformation yields a higher-dimensional representation of the lower-order, correlated model. This belief is false" (p. 4).

Two distinct classes of methods for estimating exploratory bi-factor structures have been proposed. Firstly, Jennrich and Bentler (2011, 2012) formulated the direct analytical rotations based on the quartimin and geomin rotations. Unfortunately, such procedures have been proved to be either unreliable for recovering complex structures (bi-quartimin) or prone to present local minima and factor collapse problems (bi-geomin; Hattori, Zhang, & Preacher, 2017; Mansolf & Reise, 2016). Secondly, the rediscovery of the target rotation as an alternative for recovering complex structures has prompted the emergence of bi-factor exploratory methods based on this rotation.

Automatic Specification of Target Matrices

A target rotation (also called "Procrustes rotation") is a semi-confirmatory procedure intended to approximate a rotated solution toward a pre-defined factor pattern. To do so, a target matrix (**B**), with same dimensions as the factor loading matrix (**Λ**), must be defined as follows: If the value of an entry is expected to be negligible in the targeted rotated solution, that entry is fixed as zero in **B**. Otherwise, the entry is left unspecified (partially specified target rotation; Browne, 1972). Afterward, the target rotation is performed by minimizing the following loss function:

$$f(L) = \sum_{j=1}^{m} \sum_{i \in I_j} \left(\lambda_{ij} - b_{ij}\right)^2, \qquad (1)$$

where λ_{ij} is a loading in the ith row and jth column of **Λ**, b_{ij} is an entry of the ith row and jth column of **B**, and I represents the subscript for fixed target loadings.

Although researchers can manually specify **B** according to their theoretical expectations, modern approaches rely on the automatic detection of which loadings should be fixed as zero in **B**. Reise et al. (2011) firstly advocated the use of automated bi-factor target rotations. They argued that even though solutions obtained by means of a Schmid-Leiman transformation result in biased estimates, they could provide the basis for accurately defining partially specified target matrices. This procedure led to the adequate recovery of exploratory bi-factor models, especially when cross-loadings were present in the structure.

Iteration of Partially Specified Bi-Factor Target Matrices

Moore, Reise, Depaoli, and Haviland (2015) demonstrated that an iterative target rotation procedure, where each rotated target solution formed the basis for defining a new target matrix, led to improved target specification and that such refinement directly translated into enhanced rotation accuracy. Abad et al. (2017) showed in an extensive simulation that SLi, which combines an iterative target rotation with Reise's et al. (2011) procedure, outperformed the original non-iterative method as well as both direct analytical bi-factor rotations.

[1] The term "bi-factor" is restricted hereafter to models not presenting the aforementioned constraints between general and specific factor loadings.

 © 2019 Hogrefe Publishing

${protocol}://econtent.hogrefe.com/doi/pdf/10.1027/1614-2241/a000163 - Eduardo Garcia-Garzon <eduardo.garciag@uam.es> - Saturday, July 27, 2019 12:14:13 AM - IP Address:83.34.101.73

Unfortunately, all automated target methods hitherto presented request researchers to decide a single cut-off for separating expected substantive and negligible factor loadings. While recommendations for selecting such a threshold do exist for the general exploratory factor analysis (EFA) case (between .30 and .50; Izquierdo, Olea, & Abad, 2014), the impact of this choice on rotation performance has been overlooked in the literature. Ultimately, theoretical recommendations could be misleading in a given factor analysis, as a suitable cut-off point may be dependent on the specific conditions of the factor model estimation (i.e., sample size) or could vary across factors when these differ in their factor loading distributions. Therefore, a set of cut-offs would yield a correct targeted rotated solution if, and only if, they correctly identify, for each specific factor separately, all, and exclusively, the near-zero loadings in the targeted rotated matrix (i.e., minimizes target rotation criterion; Kiers, 1994). Consequently, a single, fixed cut-off point would result in misspecification errors when the strength of the involved factors varies, and the structure departs from a simple bi-factor model. The study of such conditions (illustrated below) demonstrated that cut-offs shall be empirically estimated rather than arbitrary chosen.

Empirical Bi-Factor Target Rotation Based on Loading Differences

Under a bi-factor solution where all specific factors follow the simplest structure case, appropriate cut-off points that lead to a correct target definition might be easily found. If the model is properly estimated (i.e., adequate sample size), each specific factor yields a factor loading distribution with two distinctive groups: a first set of loadings near zero and another encompassing substantive loadings. Hence, such appropriate cut-offs will correctly fix to zero in the target matrix all near-zero loadings.

To ensure better separation between near-zero and substantive loadings, all factor loadings are traditionally transformed by means of Kaiser's normalization. This procedure consists in multiplying each $\boldsymbol{\Lambda}$ row by the inverse of the square root of the correspondent row communality. Normalization is also usually applied in this field to avoid the fact that large rows of $\boldsymbol{\Lambda}$ have more influence in the rotation criteria than smaller rows (Jennrich, 2004). Additionally, and to further increase substantive factor loading saliency, normalized loadings can be squared (Lorenzo-Seva, 1999). After applying such transformations, the size of each squared normalized factor loading becomes a function of the item complexity (i.e., number and magnitude of cross-loadings) and, for complex items, of its primariness (i.e., if the substantive loading is the primary or a secondary loading). For instance, for items following a simple population bi-factor structure, the matrix of squared row-normalized loadings should be composed by zero and one entries, which would facilitate setting a correct cut-off. A detailed discussion of the effects of such transformations can be found in Lorenzo-Seva (1999) and Browne (2001).

The presence of small cross-loadings in one specific factor, which is always to be expected (Asparouhov & Muthén, 2009; Morin et al., 2015), introduces a third group of items whose (squared normalized) factor loadings are somewhere situated between the near-zero (i.e., for items non-loading on the factor) and near-one loadings (i.e., for factorially simple items only loading in the factor). To be bear in mind that if, by definition (Equation 1), the rotation criterion only depends on the specified elements of the target matrix, wrongly fixing any potential minor cross-loading in the target matrix will result in impaired target rotation. Accordingly, an appropriate cut-off for each specific factor should result in such cross-loadings also being freed in the target matrix. Therefore, a statistical criterion to define which (squared normalized) factor loadings constitute meaningful cross-loadings is in need. Additionally, the impact of cross-loadings on cut-off estimation is ameliorated – or worsened – depending on factor loading primariness. Structures presenting low magnitude factor loadings will result in less distinctive groups of near-zero and substantive factor loadings, ultimately hampering cut-off point identification.

Thus, under conditions where cross-loadings and low average loadings factor are expected, it is implausible that a single, fixed cut-off point would correctly identify all elements in the target matrix. Any method applying fixed cut-offs (i.e., SLi) is therefore anticipated to commit many misspecification errors and to result in a biased factor solution. To overcome such limitations, a novel algorithm based on empirical, factor-specific cut-offs is here introduced: SLiD. This method, which is regarded as an improvement of the SLi algorithm as it introduces an empirical cut-off estimation within that algorithm, is likely to limit misspecification errors in target matrix definition and to result in improved factor recovery. To illustrate SLiD functioning, an analysis of a popular bi-factor structure presenting the aforementioned characteristics (Abad et al., 2017; Chen, West, & Sousa, 2006) is presented (Table 1).

The proposed SLiD method estimates factor-specific loading cut-offs based on the distribution of the differences between the sorted squared normalized factor loadings. To find this empirical cut-off, SLiD performs several steps aimed to separate near-zero loadings to be fixed in the target matrix and meaningful specific loadings to be freed (note that general factor loadings are always freed, so they are not evaluated when estimating cut-offs).

Starting from a standardized SL solution (Panel A, Table 1), the squared normalized loadings are computed (Panel B, Table 1). Secondly, squared normalized loadings are sorted

© 2019 Hogrefe Publishing

${protocol}://econtent.hogrefe.com/doi/pdf/10.1027/1614-2241/a000163 - Eduardo Garcia-Garzon <eduardo.garciag@uam.es> - Saturday, July 27, 2019 12:14:13 AM - IP Address:83.34.101.73

Table 1. SLiD algorithm first iteration target matrix definition for Chen et al. (2006) data

Panel A: Original SL solution					Panel B: Squared normalized loadings				Panel E: First iteration target matrix			
Item	SF1	SF2	SF3	SF4	SF1	SF2	SF3	SF4	SF1	SF2	SF3	SF4
1	.67	.00	.01	−.01	1.00	.00	.00	.00	NA	0	0	0
2	.47	.06	−.01	.04	0.98	.02	.00	.01	NA	0	0	0
3	.64	.01	.01	.02	1.00	.00	.00	.00	NA	0	0	0
4	.67	.00	−.02	−.03	1.00	.00	.00	.00	NA	0	0	0
5	.63	.00	.04	.00	1.00	.00	.00	.00	NA	0	0	0
6	.02	.54	.04	−.03	.00	.99	.01	.00	0	NA	0	0
7	−.01	.38	.04	.07	.00	.96	.01	.03	0	NA	0	0
8	.06	.57	−.04	.06	.01	.97	.00	.01	0	NA	0	0
9	−.03	.39	.15	−.02	.01	.86	.13	.00	0	NA	NA	0
10	.02	−.04	.36	.02	.00	.01	.98	.00	0	0	NA	0
11	.05	.00	.31	.08	.02	.00	.92	.06	0	0	NA	0
12	−.03	.06	.33	−.02	.01	.03	.96	.00	0	0	NA	0
13	.24	.01	.17	.09	.61	.00	.31	.09	NA	0	NA	0
14	.13	.05	.27	.01	.18	.03	.79	.00	NA	0	NA	0
15	.03	−.02	.00	.60	.00	.00	.00	1.00	0	0	0	NA
16	.00	.13	.02	.44	.00	.08	.00	0.92	0	0	0	NA
17	−.02	.02	.01	.52	.00	.00	.00	1.00	0	0	0	NA
Average	.21	.13	.10	.11	.34	.23	.24	.18				

Panel C: Sorted squared normalized loadings								Panel D: One-lagged differences distribution							
Item	SF1	Item	SF2	Item	SF3	Item	SF4	Item	SF1	Item	SF2	Item	SF3	Item	SF4
1	1.00	6	.99	10	.98	17	1.00	1	.00	6	.02	10	.02	17	.00
3	1.00	8	.97	12	.96	15	1.00	3	.00	8	.02	12	.04	15	.08
4	1.00	7	.96	11	.92	16	.92	4	.00	7	.09	11	.13	16	***.83***
5	1.00	9	.86	14	.79	13	.09	5	.02	9	***.78***	14	.48	13	.02
2	0.98	16	.08	13	.31	11	.06	2	.37	16	.05	13	.18	11	.03
13	.61	12	.03	9	.13	7	.03	13	.43	12	.00	9	***.12***	7	.02
14	.18	14	.03	7	.01	8	.01	14	***.16***	14	.01	7	.01	8	.00
11	.02	2	.02	6	.01	2	.01	11	.01	2	.00	6	.00	2	.00
8	.01	10	.01	8	.00	12	.00	8	.00	10	.01	8	.00	12	.00
12	.01	17	.00	5	.00	6	.00	12	.00	17	.00	5	.00	6	.00
9	.01	15	.00	16	.00	10	.00	9	.00	15	.00	16	.00	10	.00
10	.00	13	.00	4	.00	9	.00	10	.00	13	.00	4	.00	9	.00
15	.00	3	.00	2	.00	4	.00	15	.00	3	.00	2	.00	4	.00
17	.00	1	.00	17	.00	14	.00	17	.00	1	.00	17	.00	14	.00
6	.00	4	.00	3	.00	3	.00	6	.00	4	.00	3	.00	3	.00
7	.00	5	.00	1	.00	1	.00	7	.00	5	.00	1	.00	1	.00
16	.00	11	.00	15	.00	5	.00	16	–	11	–	15	–	5	–
Average									*.06*		*.06*		*.06*		*.06*

Notes. SF = specific factor (SF1: Cognition; SF2: Vitality; SF3: Mental health; SF4: Disease worry); Panels divided as item order is changed between Panels B and C. In Panel A, substantive loadings appear shadowed in strong gray and cross-loadings in light gray. First one-lagged difference above factor difference average is presented bolded and italicized.

per factor to detect a "jump" in the sequence (Panel C, Table 1). A similar strategy was deployed by Jennrich (2004) to detect the number of near-zero values to be fixed in a target matrix in a Simplimax rotation (Kiers, 1994) by visually detecting the first meaningful "jump" in a SAL plot (sorted absolute loadings plots; Figure 5, Jennrich, 2004). Thirdly, one-lagged differences are computed (differences between a factor loading and its immediate predecessor; Panel D, Table 1). For each specific factor (e.g., SF1), small one-lagged differences are expected for blocks of homogenous items. For instance, for SF1 close-to-zero one-lagged differences are found for the group of non-loading items (i.e., 11, 8, 12, 9, 10, 15, 17, 6, 7, and 16) and for almost all the group of simple items (i.e., 1, 3, 4, and 5). In contrast,

${protocol}://econtent.hogrefe.com/doi/pdf/10.1027/1614-2241/a000163 - Eduardo Garcia-Garzon <eduardo.garciag@uam.es> - Saturday, July 27, 2019 12:14:13 AM - IP Address:83.34.101.73

the higher one-lagged differences represent some "jumps" in the distribution of loadings (i.e., .16, .43, and .37). The first "jump" reveals a potential candidate for representing the separation of the last near-zero loading (i.e., .02) and the first substantive loading (i.e., .18).

We propose the average lagged difference as criterion to identify which items represent the boundary between such groups (in italics, last row in Panel D, Table 1; e.g., for SF1, .06), setting the first one-lagged difference in the sorted distribution with a value higher than the criterion as the first "jump" (bolded and italicized values in Panel D, Table 1; e.g., for SF1, .16). The rationale behind is that the number of "jumps" should be necessarily a small proportion of the one-lagged differences. Lastly, for each factor, the cut-off is settled to the factor loading corresponding to the lower endpoint of the interval represented by this difference (Panel E, Table 1; e.g., for SF1, item 11). In Table 1 example, when compared previously applied cut-offs such as .15 (Reise et al., 2011) or .20 (Abad et al., 2017), SLiD found three additional small cross-loadings to be additionally freed in the target matrix (Panel E, Table 1).

It is anticipated that the SLiD method will find, for each specific factor, an accurate cut-off regardless of the distribution of the loading differences across factors. Unfortunately, SLiD would not produce a satisfactory solution if a non-identified solution is to be produced (e.g., if at least $(j-1)/2$ targets in each column are not fixed; Asparouhov & Muthén (2009); condition C1 to C3 adapted to orthogonal structures in Peeters, 2012). Therefore, SLiD evaluates if Peeters (2012) conditions C1 to C3 are met in the target matrix (in a similar fashion to *Mplus* target rotation; Asparouhov & Muthén, 2009; Myers et al., 2013). If any of such conditions are not met, the smallest non-fixed loading of the sorted normalized factor loading distribution would be fixed in the target rotation. Due to space constrains, readers are remitted to original sources for a more detailed explanation of mentioned factor identification conditions.

To conclude, the objective of this article is three folded: first, to investigate for the first time the effect of applying arbitrary cut-offs on bi-factor target rotation, second, to understand the properties of the new proposed algorithm for exploratory bi-factor modeling (SLiD), and third, to compare the performance of SLiD and SLi under a set of realistic conditions by means of a Monte Carlo simulation.

Method

Simulation Design

A Monte Carlo study was designed to compare SLiD against the SLi target rotation using arbitrary cut-offs. For SLi, four cut-off points were examined (.05, .10, .15, and .20), covering the range previously considered in bi-factor exploratory research (Abad et al., 2017; Reise, Moore, & Haviland, 2010; Waller, 2017). Each rotation is hereafter referred to as SLi followed by the applied cut-off point.

Table 2. Manipulated factors included in the Monte Carlo study

Factor	Levels
Sample size (N)	Low (n = 500); Medium (n = 1,000); High (n = 2,000).
Indicators per specific factor (VAR.SF)	Low (4); Medium (6); High (12).
Number of specific factors (NUM.SF)	Low (3); High (6).
Cross-loading average size (CROSS.SD)	Low (M = 0.00, SD = 0.05); Medium (M = 0.00, SD = 0.10); High (M = 0.00, SD = 0.15).

Several variables were manipulated in the following Monte Carlo simulation (Table 2): sample size (N), number of variables per specific factor (VAR.SF), number of specific factors (NUM.SF), and cross-loading size (CROSS.SD), yielding $3 \times 3 \times 2 \times 3 = 54$ fully crossed conditions. Population bi-factor structures were simulated as follows: (a) General factor loadings ranged from .575 to .625 in equal increments and were randomly sorted and (b) for each structure, specific factors were simulated varying in the average loading size (high, medium, and low); specifically, loadings ranged in equal increments from .575 to .625 (first or first two factors, depending upon NUM.SF), from .425 to .475 (central or two central factors), and from .275 to .325 (last or last two factors).

For each structure, the first two items in each factor were set as markers (all their cross-loadings were fixed to zero). For the remaining indicators, cross-loadings were drawn from a normal distribution with a mean of 0 and a standard deviation of .05 (low condition: 95% of cross-loadings ranged between ±.098), .10 (medium condition: 95% of cross-loadings ranged between ±.196), and .15 (high condition: 95% of cross-loadings ranged between ±.294). This procedure, also found in Meade (2008), simulated cross-loadings distributions that are consistent with those commonly found in the literature (Bollmann, Heene, & Küchenhoff, 2015). If either the population correlation matrix was not positive semi-definite (i.e., minimum eigenvalue ≤ 0) or the item communalities were higher than .90, all simulated cross-loadings were replaced. Simulated average item communalities across conditions depended upon number of factors simulated and cross-loading size, ranging from .580 (i.e., NUM.SF = 3; CROSS.SD = .05) to .732 (i.e., NUM.SF = 6; CROSS.SD = .15).

© 2019 Hogrefe Publishing

${protocol}://econtent.hogrefe.com/doi/pdf/10.1027/1614-2241/a000163 - Eduardo Garcia-Garzon <eduardo.garciag@uam.es> - Saturday, July 27, 2019 12:14:13 AM - IP Address:83.34.101.73

Data Generation

A total of 800 sample data matrices were simulated for each condition. First, for each replication, the cross-loadings were generated. Then, the population correlation matrix, obtained by inserting unities in the diagonal of the reproduced correlation matrix, was used to generate a matrix of random standard normal observed variables with dimensions of sample size (*N*) per number of specific variables (VAR.SF) with the function *rmvnorm* (mvnorm package; Gentz et al., 2017; R Development Core Team, 2017).

Rotation Methods

Scripts in R (R Development Core Team, 2017) were applied to obtain the rotated solutions. The SLi solutions were computed using the SLi function in Abad et al. (2017), while the SLiD code is provided in the Electronic Supplementary Material, ESM 1. In all the cases, the unweighted least squares estimator was used for factor extraction (psych package; Revelle, 2017).

Accuracy Criteria

Accuracy was assessed by means of Tucker's congruence coefficient (cc; Tucker, 1951), which measures similarity between two vectors. The cc values are bounded between −1 and 1, where values ranging from .85 to .94 are considered as *fair* and values over .95 reflect that the two factors should be considered equal (Lorenzo-Seva & ten Berge, 2006).

To accurately compute cc values, estimated factors must be accordingly aligned with their correspondent population factor. Firstly, an estimated factor was reverted if the sum of factor loadings was negative. Secondly, cc was computed for every possible factor permutation, retaining the solution that minimized the average deviation between that solution and the population matrix.

To compare cc means across conditions, analyses of variance were performed using the jamovi program 0.8.1.13 (jamovi Project, 2018). Partial eta squared (η_p^2) effect sizes were reported, where $\eta_p^2 > .01$, $\eta_p^2 > .06$, and $\eta_p^2 > .14$ were considered as small, medium, and large effects, respectively (Cohen, 1988). Additionally, to depict how each distinctive factor was recovered (i.e., general factor and specific factors with high, medium, and low average factor loadings), the percentage of samples across conditions showing fair (i.e., cc $\geq$ .85) and good (i.e., cc $\geq$ .95) recovery were computed for the different rotation methods.

Results

Mean cc for each rotation procedure and simulated condition are depicted in Table 3. Firstly, SLiD was the preferred method in all conditions, consistently producing the highest cc for all levels considered. Secondly, SLi with a fixed cut-off point showed an inadequate overall performance, with mean cc values decreasing as cut-offs increased. In this line, it is noteworthy that SLi.20 only reached a fair recovery (cc $\geq$ .85) for one marginal condition (cc = .887 with low cross-loadings). In general, for all methods increasing either the sample size or the number of variables per specific factor resulted in increased factor recovery (except for SLi.20), while increasing the number of factors or the cross-loading size decreased their accuracy.

Table 4 provides the analysis of variance (ANOVA) effect sizes for the SLiD and SLi with arbitrary cut-off point methods. All methods presented a medium-to-large two-way interaction involving Cross-Loading Size × Number of Factors (NUM.SF × CROSS.SD), implying that the negative effect of cross-loadings was reduced in small structures (e.g., cc decrease for NUM.SF = 3; SLi.05 = .12; cc decrease for NUM.SF = 6; SLi.05 = .23). The interaction effect size diminished as the cut-off for SLi increased (SLi.05 η_p^2 = .307; SLi.10 η_p^2 = .290; SLi.15 η_p^2 = .251; SLi.20 η_p^2 = .209), being SLiD the most robust method (SLiD η_p^2 = .066). Lastly, no method was strongly affected by either sample size ($.007 \leq \eta_p^2 \leq .020$) or the number of variables per factor ($.003 \leq \eta_p^2 \leq .098$).

Previous results were concerned with the mean cc values for each complete structure (i.e., averaging over factors of different average loading size), which could provide a limited view of the performance of the rotation methods. When analyzing the percentage of replicates that were satisfactorily recovered across conditions (Table 5, upper panel), SLiD performed better than SLi with any fixed cut-off when recovering all types of factors. The largest discrepancies were found for cc > .95 with factors composed of high loadings ($\bar{\lambda}$ = .60; difference ≈ 40%). Nevertheless, results illustrated that not all factors were equally recovered and that each factor appropriate cut-off was dependent upon factor loading magnitude. Regarding general, high, and medium size factors, all arbitrary cut-offs performed similarly well (average cc difference = .001), where increasing the cut-off point somewhat improved the recovery for these factors.

However, the opposite effect was found when analyzing the factors with low magnitude loadings. As the cut-off increased, the SLi performance was greatly diminished (cc differences: SLi.05 and SL.10 = .049; SLi.05 and SL.15 = .076; SLi.05 and SL.20 = .212). It is noteworthy that even though SLiD still outperformed all SLi methods for

 © 2019 Hogrefe Publishing

${protocol}://econtent.hogrefe.com/doi/pdf/10.1027/1614-2241/a000163 - Eduardo Garcia-Garzon <eduardo.garciag@uam.es> - Saturday, July 27, 2019 12:14:13 AM - IP Address:83.34.101.73

Table 3. Marginal mean factor recovery for the SLi with arbitrary cut-off points and SLiD algorithms

Variable/Level	SLiD	SLi.05	SLi.10	SLi.15	SLi.20
N					
500	***.920***	.845	.842	.835	.817
1,000	***.927***	.851	.849	.843	.821
2,000	***.930***	.858	.857	.851	.829
VAR.SF					
Low (4)	***.911***	.834	.834	.834	.827
Medium (8)	***.930***	.855	.851	.844	.822
High (12)	***.936***	.866	.861	.851	.819
NUM.SF					
Low (3)	.930	.886	.881	.868	.835
High (6)	***.921***	.817	.817	.818	.810
CROSS.SD					
Low (.05)	***.965***	.934	.931	.921	.887
Medium (.10)	***.932***	.858	.855	.851	.831
High (.15)	***.880***	.762	.760	.757	.749
Average	***.926***	.851	.849	.843	.826

Notes. *N* = sample size; VAR.SF = variables per specific factor; NUM.SF = number of specific factors; CROSS.SD = standard deviation for cross-loading simulation. Congruence coefficients ≥ .85 appear shadowed in gray. Best cc for each condition appears bolded and italicized.

Table 4. Univariate analysis of variance (ANOVA) effect sizes for SLi with arbitrary cut-off points and SLiD

Effect type/Variables	SLiD	SLi.05	SLi.10	SLi.15	SLi.20
Main effects					
N	.014	.020	.020	.017	.007
VAR.SF	.095	.098	.061	.018	.003
NUM.SF	.051	.538	.462	.297	.102
CROSS.SD	.519	.771	.737	.646	.454
Two-way interactions					
NUM.SF × CROSS.SD	.066	.307	.290	.251	.209

Notes. *N* = sample size; VAR.SF = number of variables per specific factor. NUM.SF = number of specific factors; CROSS.SD = cross-loading standard deviation. The dependent variable in the ANOVAs was the congruence coefficient. Only medium (η_p^2 > .06) or larger effects are presented, with large (η_p^2 > .14) effects shadowed in gray.

factors that had low loadings, it still struggled to adequately recover these factors (SLiD low cc = .829).

Despite the overall results, it should be noted that SLi with a fixed cut-off point provides a good recovery under optimal conditions (i.e., CROSS.SD = .05; Table 5, lower panel). All SLi methods with arbitrary cut-offs satisfactorily recovered the general, the high, and the medium loading size factors (i.e., all average cc ≥ .85), benefitting again from applying higher cut-offs. Once more, SLi methods with arbitrary cut-offs failed to recover the low loadings factors, especially when applying higher cut-offs (SLi.15 cc low loadings = .775; SLi.20 cc low loadings = .641).

Discussion

Almost 80 years after their appearance, bi-factor models have become a major research topic in the psychometric literature. This article presented, for the first time, evidence of the consequences of using arbitrary cut-off points for exploratory bi-factor target rotation. Furthermore, a new algorithm for the empirical estimation of accurate cut-offs under realistic conditions was introduced in this study. This new proposal was evaluated by means of an extensive Monte Carlo simulation against alternative methods applying fixed cut-offs.

Main Findings

The Use of Fixed Cut-Off Points

Recommendations found in the EFA literature (i.e., .30–.50; Izquierdo et al., 2014) were not satisfactorily translated into exploratory bi-factor analysis. Due to item variance partition specified in the bi-factor model, specific factor loadings could be expected to be lower than in other EFA applications, entailing the necessity of adapting factor saliency thresholds. According to our expectations, lower cut-offs benefited factor recovery when applying bi-factor target rotation. However, it was demonstrated that the impact of the arbitrary cut-off point was dependent upon the factor loading distribution, where higher cut-offs were significantly more affected by the presence of a factor with lower substantive loadings. Remarkably, factors with either high- or medium-sized loadings were recovered with similar accuracy by all fixed cut-offs evaluated and with a precision similar to previous studies (if similar conditions were evaluated; Table 5, lower panel). Furthermore, the inclusion of a low average factor loadings in combination with the presence of several cross-loadings caused a substantial decrease in the overall accuracy for SLi with higher cut-off points (Table 5), particularly if compared with its performance reported in a previous study (Abad et al., 2017). Lastly, it should be acknowledged that, as also reported in earlier research, simple bi-factor structures, with none to a few low magnitude cross-loadings and high to medium average loadings factors, are expected to be correctly recovered by all methods, including SLi with high cut-off points (Abad et al., 2017; Table 2). Nevertheless, in realistic settings when structures are expected to either present specific factors with low average factor loadings, items with strong cross-loadings, or its combination, SLiD should be preferred.

Empirical Estimation of Cut-Off Points

The proposed method based on loading differences (SLiD), which was used in conjunction with SLi rotation,

${protocol}://econtent.hogrefe.com/doi/pdf/10.1027/1614-2241/a000163 - Eduardo Garcia-Garzon <eduardo.garciag@uam.es> - Saturday, July 27, 2019 12:14:13 AM - IP Address:83.34.101.73

Table 5. Proportion of replicates with mean cc higher than .85 and .95 for SLiD and SLi across all conditions (upper panel) and for low cross-loading condition (lower panel)

Factor	Mean cc					cc > .85					cc > .95				
	SLiD	SLi.05	SLi.10	SLi.15	SLi.20	SLiD (%)	SLi.05 (%)	SLi.10 (%)	SLi.15 (%)	SLi.20 (%)	SLiD (%)	SLi.05 (%)	SLi.10 (%)	SLi.15 (%)	SLi.20 (%)
All CROSS.SD conditions															
General	.991	.990	.990	.991	.992	100	100	100	100	100	99	99	99	99	99
High	.966	.893	.896	.898	.900	99	76	77	79	79	82	42	45	47	47
Medium	.921	***.827***	***.828***	***.827***	***.824***	89	53	54	55	55	47	17	18	19	19
Low	***.829***	***.740***	***.691***	***.664***	***.528***	54	28	27	25	18	13	3	3	1	1
CROSS.SD = .05															
General	1.00	.998	.999	.998	.998	100	100	100	100	100	100	100	100	100	100
High	.990	.969	.972	.973	.973	100	100	100	100	100	100	95	98	99	99
Medium	.970	.936	.939	.940	.940	98	97	97	98	97	95	44	48	49	50
Low	.901	***.836***	***.812***	***.775***	***.641***	88	63	62	57	40	31	8	7	7	7

Notes. SLiD: Schmid-Leiman Iterative Bi-factor Difference-based Target Rotation. In case of NUM.SF=6 structures, high, medium, and low loadings factors represent the average of the two factors with such characteristics, respectively. Mean cc below .85 is presented bolded and italicized. Conditions with cells with less than 50% appear shadowed in gray.

was defined as an iterative, factor-specific, empirically estimated, automatic algorithm for defining partially specified target rotations based on an initial SL solution. SLiD objective is to find the first cross-loading in the sorted squared normalized factor loading distribution and to fix values below such value in the target matrix. Additionally, and to the authors' knowledge, this is the first target rotation algorithm deploying a criterion based on sorted, one-lagged differences obtained from the empirical factor loading distribution.

Results evidenced that when compared to SLi with arbitrary cut-offs, this method improved factor recovery for all the studied conditions. Far from perfect, the performance of SLiD was greatly diminished under the presence of high size cross-loadings, especially when recovering factors with low substantive loadings (Table 5). However, such a result may be due to the SLi algorithm itself (and not related to the method for defining the target matrix), as explained next.

The Pitfall of the SLi Algorithm

When inspecting replicates that were not adequately recovered, it was observed that factor collapse occurred in several cases and that the higher the magnitude of these cross-loadings, the more aggravated the problem was. Indeed, factor collapse could be deemed as the consequence of a non-uniqueness problem of the simulated bi-factor structures (Green & Yang, 2018). If by chance, or due the presence of cross-loadings (combined with the presence of factors differing in their average factor loadings), the reproduced correlation matrix presents a true dimensionality approximated to the one of a hierarchical solution (dimensionality equals the number of specific factors, not the number of specific factors + one as in the bi-factor case), SLi solutions could result in factor-collapsed solutions (which would be in accordance with the true dimensionality; Mansolf & Reise, 2016). Additionally, these collapsed solutions could also be consistent with collapsed population, non-simulated, bi-factor structures that also approximate such reproduced correlation matrices. Unfortunately, no trustworthy method for evaluating the uniqueness of a simulated bi-factor structure or to preclude factor collapse on the basis of a simulated reproduced correlation matrix prior to factor rotation has been developed yet.

Furthermore, all SLi-based algorithms apply a partially specified target rotation where all general factor loadings, and only some of the specific factor loadings, are freed in the target matrix. This leads to the general factor to absorb as much variance as possible. Therefore, if researchers aim to obtain solutions not presenting collapse, applying an SL transformation would be a viable option in some cases (i.e., as the SL solution would provide a good fit to the data). Alternatively, unreported analyses confirmed that factor collapse could be prevented in some specific cases by applying a totally specified target rotation in SLiD. Such target matrix additionally fixes a value (typically ±1) for expected substantive values in the target matrix (Browne, 2001), so that all elements in the target rotation are given a target value. However, there is no free lunch in target rotation: In a totally specified target rotation, two kinds of errors could simultaneously occur, as factor loadings are either being shrunk to zero or enlarged toward ±1. The latter cannot occur in partially specified matrices, as the only distances to be minimized are those with respect to targets of zero. Unfortunately, the effect of applying each

 © 2019 Hogrefe Publishing

${protocol}://econtent.hogrefe.com/doi/pdf/10.1027/1614-2241/a000163 - Eduardo Garcia-Garzon <eduardo.garciag@uam.es> - Saturday, July 27, 2019 12:14:13 AM - IP Address:83.34.101.73

type of target rotation to bi-factor structures remains uninvestigated.

Limitations and Future Directions

Firstly, the current findings are limited to the conditions studied in the simulation, where the aim of the selected conditions was to reproduce realistic models covering a wide range of complex structures. Secondly, neither the impact of over and under factor extraction, nor items loading only in either the general or the specific factors were considered. Lastly, only bi-factor structures with orthogonal specific factors were simulated. Though alternative bi-factor structures with oblique specific factors can be fitted, their interpretation is still a matter of controversy.

Interest in both bi-factor structures and target rotation methods has dramatically increased over the last decade. Today, the distinctions and equivalences between exploratory hierarchical and exploratory bi-factor structures have been substantively clarified (Mansolf & Reise, 2016; Waller, 2017). Additionally, while the target rotation has become one of the most compelling alternatives for studying complex structures, its generalized application is still upheld by some unresolved limitations: (a) how to identify the best number of elements to be fixed in the target matrix. The comparison of SLiD with alternative methods for objective empirical cut-off estimation (such as Promin; Lorenzo-Seva, 1999) shall be a priority of future research efforts; (b) once the correct number of elements to be fixed in **B** is known, how to deal with optimal solutions prone to local minima problems (as occurs in Simplimax); and (c) how to guarantee unique bi-factor structures and identified rotated solutions. While factor identification has received less attention than other topics in the literature, it should not be disregarded when applying target rotations.

Additionally, to ensure the predictive validity of an empirical procedure such as SLiD, a split-half cross-validation, similar to the one found in FACTOR (Ferrando & Lorenzo-Seva, 2017), can be applied as follows[2]: (a) firstly, divide the sample into two random subsets; (b) apply SLiD to the first sample to obtain an empirical, iteratively refined target matrix; (c) apply this target in a non-iterative fashion to an unrotated solution found in the second subset; (d) if step b and step c solutions are similar, researchers can be more confident regarding the validity of the target applied by SLiD. Researchers should ensure, however, that both subsets are of sufficient sample size for both solutions to be properly estimated. Moreover, if researchers are interested in assessing the accuracy of such a procedure, a bootstrapping strategy could be of employed (Raykov & Little, 1999).

Lastly, alternative methods based on target rotations have been recently proposed (BIFAD; Waller, 2017). However, the BIFAD method, which applies a totally specified target rotation, is also based on using arbitrary cut-off points. Future research should address how a loadings difference method for defining a totally specified target in the BIFAD algorithm would perform and compare it with the recommendations for defining arbitrary cut-off points, which the authors are currently investigating (Waller, 2017, p. 12).

Electronic Supplementary Material

The electronic supplementary material is available with the online version of the article at https://doi.org/10.1027/1614-2241/a000163

ESM 1. Code (.RAR)
This R Script file contains the SLiD algorithm code, as well as the necessary auxiliary functions to run the SLi function.

References

Abad, F. J., Garcia-Garzon, E., Garrido, L. E., & Barrada, J. R. (2017). Iteration of partially specified target matrices: Application to the bi-factor case. *Multivariate Behavioral Research, 52*, 416–429. https://doi.org/10.1080/00273171.2017.1301244

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 16*, 397–438. https://doi.org/10.1080/10705510903008204

Bollmann, S., Heene, M., & Küchenhoff, H. (2015). *What can the Real World do for simulation studies? A comparison of exploratory methods*. (Technical Report 181). Retrieved from Department of Statistics, University of Munich https://epub.ub.uni-muenchen.de/24518/

Browne, M. W. (1972). Orthogonal rotation to a partially specified target. *British Journal of Mathematical and Statistical Psychology, 25*, 115–120. https://doi.org/10.1111/j.2044-8317.1972.tb00482.x

Browne, M. W. (2001). The latent structure of memory: A confirmatory factor-analytic study of memory distinctions. *Multivariate Behavioral Research, 36*, 29–51. https://doi.org/10.1207/S15327906MBR3601

Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bi-factor and second order models of quality-of-life. *Multivariate Behavioral Research, 41*, 189–225. https://doi.org/10.1207/s15327906mbr4102_5

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Ferrando, P. J., & Lorenzo-Seva, U. (2017). Program FACTOR at 10: Origins, development and future directions. *Psicothema, 29*, 236–240. https://doi.org/10.7334/psicothema2016.304

[2] We are indebted to an anonymous reviewer for suggesting this cross-validation procedure.

© 2019 Hogrefe Publishing

${protocol}://econtent.hogrefe.com/doi/pdf/10.1027/1614-2241/a000163 - Eduardo Garcia-Garzon <eduardo.garciag@uam.es> - Saturday, July 27, 2019 12:14:13 AM - IP Address:83.34.101.73

Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F., Bornkamp, B., ... Hothorn, T. (2017). *mvtnorm: Multivariate Normal and t Distributions*. R package version 1.0-6. Retrieved from https://CRAN.R-project.org/package=mvtnorm

Green, S., & Yang, Y. (2018). Empirical underidentification with the bifactor model: A case study. *Educational and Psychological Measurement, 78*, 717–736. https://doi.org/10.1177/0013164417719947

Hattori, M., Zhang, G., & Preacher, K. J. (2017). Multiple local solutions and geomin rotation. *Multivariate Behavioral Research, 52*, 720–731. https://doi.org/10.1080/00273171.2017.1361312

Holzinger, K. J., & Swineford, F. (1937). The bi-factor method. *Psychometrika, 2*, 41–54. https://doi.org/10.1007/BF02287965

Izquierdo, I., Olea, J., & Abad, F. J. (2014). Exploratory factor analysis in validation studies: Uses and recommendations. *Psicothema, 26*, 395–400. https://doi.org/10.7334/psicothema2013.349

jamovi Project. (2018). *jamovi*. (Version 0.8.1.13) [Computer Software]. Retrieved from https://www.jamovi.org/

Jennrich, R. I. (2004). Rotation to simple loadings using component loss functions: The orthogonal case. *Psychometrika, 69*, 257–273. https://doi.org/10.1007/S11336-003-1136-B

Jennrich, R. I., & Bentler, P. M. (2011). Exploratory bi-factor analysis. *Psychometrika, 76*, 537–549. https://doi.org/10.1007/s11336-011-9218-4

Jennrich, R. I., & Bentler, P. M. (2012). Exploratory bi-factor analysis: The oblique case. *Psychometrika, 77*, 442–454. https://doi.org/10.1007/s11336-012-9269-1

Kiers, H. A. L. (1994). Simplimax: Oblique rotation to an optimal target with simple structure. *Psychometrika, 59*, 567–579. https://doi.org/10.1007/BF02294392

Lorenzo-Seva, U. (1999). Promin: A method for oblique factor rotation. *Multivariate Behavioral Research, 34*, 347–365. https://doi.org/10.1207/S15327906MBR3403_3

Lorenzo-Seva, U., & ten Berge, J. M. F. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology, 2*, 57–64. https://doi.org/10.1027/1614-2241.2.2.57

Mansolf, M., & Reise, S. P. (2016). Exploratory bifactor analysis: The Schmid-Leiman orthogonalization and Jennrich-Bentler analytic rotations. *Multivariate Behavioral Research, 51*, 698–717. https://doi.org/10.1080/00273171.2016.1215898

Meade, A. W. (2008, April). *Power of AFIs to detect CFA model misfit*. Paper presented at the 23rd Annual Conference of the Society for Industrial and Organizational Psychology, San Francisco, CA

Moore, T. M., Reise, S. P., Depaoli, S., & Haviland, M. G. (2015). Iteration of partially specified target matrices: Applications in exploratory and Bayesian confirmatory factor analysis. *Multivariate Behavioral Research, 50*, 149–161. https://doi.org/10.1080/00273171.2014.973990

Morin, A. J. S., Arens, A. K., & Marsh, H. W. (2015). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling: A Multidisciplinary Journal, 23*, 116–139. https://doi.org/10.1080/10705511.2014.961800

Myers, N. D., Ahn, S., & Jin, Y. (2013). Rotation to a partially specified target matrix in exploratory factor analysis: How many targets? *Structural Equation Modeling: A Multidisciplinary Journal, 20*, 131–147. https://doi.org/10.1080/10705511.2013.742399

Myers, N. D., Jin, Y., Ahn, S., Celimli, S., & Zopluoglu, C. (2015). Rotation to a partially specified target matrix in exploratory factor analysis in practice. *Behavior Research Methods, 47*, 494–505. https://doi.org/10.3758/s13428-014-0486-7

Peeters, C. F. W. (2012). Rotatitional uniqueness conditions under oblique factor correlation metric. *Psychometrika, 77*, 288–292. https://doi.org/10.1007/S11336-012-9259-3

Raykov, T., & Little, T. D. (1999). A note on procrustean rotation in exploratory factor analysis: A computer intensive approach to goodness-of-fit evaluation. *Educational and Psychological Measurement, 59*, 47–57. https://doi.org/10.1177/00131649921969730

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47*, 667–696. https://doi.org/10.1080/00273171.2012.715555

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment, 92*, 544–559. https://doi.org/10.1080/00223891.2010.496477.Bifactor

Reise, S. P., Moore, T., & Maydeu-Olivares, A. (2011). Target rotations and assessing the impact of model violations on the parameters of unidimensional item response theory models. *Educational and Psychological Measurement, 71*, 684–711. https://doi.org/10.1177/0013164410378690

Revelle, W. (2017). *Psych: Procedures for personality and psychological research*. Northwestern University: Evanston, IL. (Version 1.7.8). Retrieved from https://CRAN.R-project.org/package =psch

R Development Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Schmid, J., & Leiman, J.M. (1957). The development of hierarchical factor solutions. *Psychometrika, 22*, 53–61. https://doi.org/10.1007/BF02289209

Trendafilov, N. T. (1994). A simple method for procrustean rotation in factor analysis using majorization theory. *Multivariate Behavioral Research, 29*, 385–408. https://doi.org/10.1207/s15327906mbr2904_4

Tucker, L. R. (1951). *A method for synthesis of factor analysis studies* (Personal Research Section Report No. 984) . Washington, DC: Department of the Army

Waller, N. G. (2017). Direct Schmid-Leiman transformations and rank-deficient loadings matrices. *Psychometrika, 83*, 858–870. https://doi.org/10.1007/s11336-017-9599-0

History
Received March 21, 2018
Revision received September 4, 2018
Accepted October 29, 2018
Published online May 27, 2019

Acknowledgments
The authors declare no acknowledgments.

Conflict of Interest
The authors declare no conflict of interest.

Funding
This research was partially supported jointly by the Spanish Ministry of Education, Culture and Sports via a doctoral grant to the first author (FPU15/03246), by grant PSI2013-44300-P (Ministerio de Economia y Competitividad, Spain) and Fondo Social Europeo and grant PSI2017-85022-P (Ministerio de Economia y Competitividad, Spain) and Fondo Social Europeo to the second author.

ORCID
Eduardo Garcia-Garzon
https://orcid.org/0000-0001-5258-232X

 © 2019 Hogrefe Publishing

Eduardo Garcia-Garzon
Departamento de Psicología Social y Metodología
Facultad de Psicología
Universidad Autónoma de Madrid
Ciudad Universitaria de Cantoblanco
28049 Madrid
Spain
eduardo.garciag@uam.es

Eduardo Garcia-Garzon is a PhD candidate at the Department of Social Psychology and Methodology at the Universidad Autónoma de Madrid, Spain. His research interests are related to bi-factor modeling from an exploratory approach, factor rotation, and multidimensional assessment.

Francisco José Abad is Associate Professor of Psychometrics in the Department of Social Psychology and Methodology at the Universidad Autónoma de Madrid, Spain. His research interests include computerized adaptive testing, item response theory, factor analysis, measurement of Intelligence, as well as the application of multidimensional psychometric models for personality measurement.

Luis Eduardo Garrido is a Professor of Psychometrics in the Department of Psychology at Pontificia Universidad Católica Madre y Maestra, Dominican Republic. His research interests include the dimensionality assessment of categorical variables, structural equation modeling, mixture modeling, assessment of model fit, scale development and validation, and wording bias in self-report measures.

Appendix

Table A1. Algorithm 1: Schmid-Leiman Iterative Bi-factor Difference-based Target Rotation (SLiD)

1. Start compute a Schmid-Leiman solution.
2. repeat
3. **for** factor loading matrix.
4. **if** factor column is a specific factor.
5. define target rotation matrix based on sorted squared normalized factor loading differences
6. **if** conditions C1 to C3 defined in Peeters (2012) are met:
the smallest non-fixed element of sorted normalized factor loading vector is fixed in the target matrix
7. **else** all entries are freed in the target rotation matrix.
8. **end if**
9. perform partially specified target rotation using gradient projection algorithm.
10. end for
11. until convergence criterion achieved (equal target matrix in two consecutive solutions).

${protocol}://econtent.hogrefe.com/doi/pdf/10.1027/1614-2241/a000163 - Eduardo Garcia-Garzon <eduardo.garciag@uam.es> - Saturday, July 27, 2019 12:14:13 AM - IP Address:83.34.101.73

© 2019 Hogrefe Publishing

Chapter 4

Searching for G: A New Evaluation of SPM-LS Dimensionality

This study was published in the journal *The Journal of Intelligence*. The following pages include this publication. This article was first published in June 28, 2019. This article was awarded the 2019 Travel Award of the Journal of Intelligence to attend the 90th International Meeting of the Psychometric Society (Santiago de Chile, July, 2019). The article can be accessed at https://doi.org/10.3390/jintelligence7030014

Journal of *Intelligence*

MDPI

Article

Searching for G: A New Evaluation of SPM-LS Dimensionality

Eduardo Garcia-Garzon [1,*], Francisco J. Abad [1] and Luis E. Garrido [2]

[1] Facultad de Psicología, Universidad Autónoma de Madrid, 28049 Madrid, Spain
[2] Facultad de Psicología, Pontificia Universidad Católica Madre y Maestra, Santo Domingo 10109, Dominican Republic
* Correspondence: eduardo.garciag@uam.es; Tel.: +34-91-497-8750

Received: 26 April 2019; Accepted: 25 June 2019; Published: 28 June 2019

Abstract: There has been increased interest in assessing the quality and usefulness of short versions of the Raven's Progressive Matrices. A recent proposal, composed of the last twelve matrices of the Standard Progressive Matrices (SPM-LS), has been depicted as a valid measure of *g*. Nonetheless, the results provided in the initial validation questioned the assumption of essential unidimensionality for SPM-LS scores. We tested this hypothesis through two different statistical techniques. Firstly, we applied exploratory graph analysis to assess SPM-LS dimensionality. Secondly, exploratory bi-factor modelling was employed to understand the extent that potential specific factors represent significant sources of variance after a general factor has been considered. Results evidenced that if modelled appropriately, SPM-LS scores are essentially unidimensional, and that constitute a reliable measure of *g*. However, an additional specific factor was systematically identified for the last six items of the test. The implications of such findings for future work on the SPM-LS are discussed.

Keywords: Raven matrices; Standard Progressive Matrices test; dimensionality; bi-factor; parallel analysis; target rotation; exploratory graph analysis

1. Introduction

The Standard Progressive Matrices (i.e., SPM [1]), in any of its forms, constitutes one of the most applied tests for measuring general intelligence (*g*). Due to its considerable length (60 items), there has been a growing interest in developing short versions of this test. Unfortunately, the available short versions—such as the Advanced Progressive Matrices tests (i.e., APM)—present substantial shortcomings [2]. Consequently, [2] proposed the SPM-LS, a new short version of the SPM test based on its last, most-difficult 12 matrices of this test. These items consist of non-verbal stimuli where each item presents a single correct answer and seven distractors. In its recent validation, the SPM-LS scores were analysed using exploratory and confirmatory factor analyses as well as item response theory models as follows: After concluding that the SPM-LS scores were sufficiently unidimensional, individual responses were modelled with the 1 to 4 parameter logistic models. Additionally, a three-parameter nested logistic model was applied to recover relevant information from responses to the different distractors. Remarkably, the original authors concluded that the SPM-LS was a superior alternative to the APM test ([2]; p.113), and encouraged other researchers to re-analyse this dataset by making it publicly available and by opening a call for papers on the matter in the Journal of Intelligence.

As part of this call, this investigation will re-evaluate [2] claim of SPM-LS being essentially unidimensional. This claim is vital to understand if SPM-LS represents a valid measure of *g* and represent a necessary assumption for many of the following analysis presented by the original authors. As [2] acknowledged that "SPM-LS may not be a purely unidimensional measure" (p.114), we decided to analyse SPM-LS dimensionality by expanding the original approaches with the application of network-based exploratory analysis and bi-factor modelling.

1.1. On the Progressive Matrices Dimensionality

Few consensuses are more extended in the intelligence literature than the belief that the SPM test [1] represents a consistent measure of general intelligence (*g*; Panel A, Figure 1). Even though this claim has received overwhelming support in the literature [3–5], other authors have considered general intelligence to be a broader construct to be measured with different tasks and item formats [6]. Be that as it may, support for strict unidimensionality has historically been equivocal for short SMP versions such as the APM test. As early as 1981, some authors found evidence of an orthogonal two-factor model [7,8] were among the first authors to suggest that a nuisance factor, corresponding to a "speed factor", could be found for APM scores (Panel C, Figure 1). [3] found that the two-factor proposed in [2] fitted the data better than the single factor model if the inter-factor correlation was estimated. Nevertheless, the high magnitude of this correlation (i.e., 0.89; Panel B, Figure 1; [3]), in conjunction with the inspection of fit statistics, was taken as evidence in favour of a unidimensional model. Since then, other authors on the field have supported [3] conclusions [4,5].

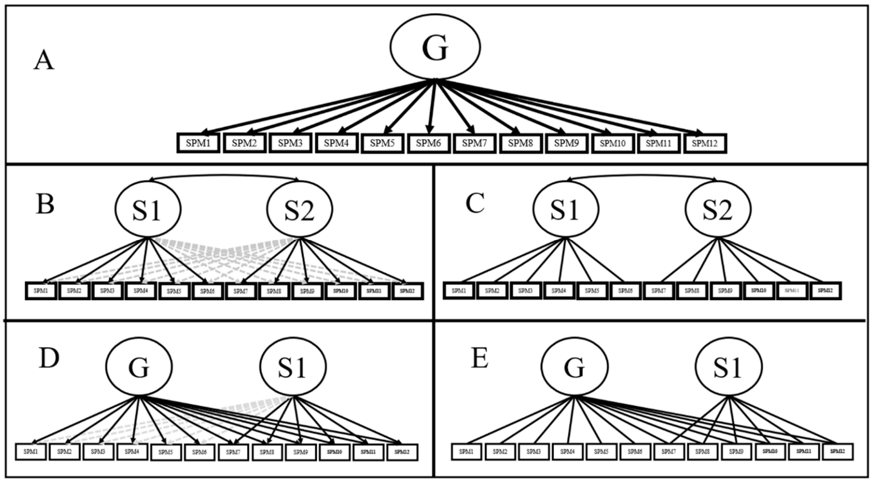

Figure 1. Schematic representation of theoretical SPM-LS models: (**A**): Unidimensional model; (**B**) Exploratory bi-dimensional model; (**C**): Confirmatory bi-dimensional model; (**D**): Exploratory bi-factor model; (**E**): Confirmatory bi-factor model. Arrows in black represent estimated paths for CFA models, and untargeted loadings in EFA models. Grey arrows represent targeted (minimised) loadings during EFA target rotation.

Recent applications of bi-factor modelling offered new insights regarding the dimensionality of the APM, as well as the role of potential secondary factors (Panel E, Figure 1). As the bi-factor model simultaneously estimates a general plus several orthogonal specific factors [9], it provides a clear separation of such different sources of variation. Noteworthy, as specific factors only account for a variance that is residual to the general factor [10], the bi-factor model can shed light about APM scores being affected by other sources of variation in addition to *g*. Indeed, APM scores do not represent a perfect measure of *g* and that alternative tests (such as Arithmetic Applications from the Weschler Adult Intelligence Scale included in the Minnesota Study of Twins Reared Apart [11]) were more strongly loaded by *g* in some specific datasets [12]. Moreover, approximately 50% of the APM true variance could be related to *g*, with 10% belonging to specific factors, and as much as 25% related to test specific variance [12]. Confirmatory bi-factor models (i.e., BCFA) also presented a better fit to the data than the unidimensional model in alternative applications such as the Coloured Progressive Matrices test (an adaptation of the APM test to children from five to 11 years old; [13]).

Most recently, the presence of additional dimensions accounting for speed factors (as well as other effects such as item position) in APM scores [14] has been linked to specific learning types [15] as well as developmental differences [16]. In either case, such evidence reflects these factors possibly being of theoretical interest. Nevertheless, the presence and nature of these additional factors in APM scores is still a matter of contention.

1.2. Modern Approaches Towards Dimensionality Assessment

Most authors have generally based their decisions regarding the unidimensionality of the SPM scores either by applying eigenvalue-based dimensionality assessment methods (i.e., parallel analysis), by comparing fit statistics from CFA models (i.e., comparing the Comparative Fit Index) or by inspecting general factor reliability (i.e., Cronbach's α). Unfortunately, these three strategies have substantial shortcomings: Firstly, parallel analysis could hide relevant sources of variation while overestimating the presence of a single factor [17]. Also, its estimation is substantially affected by the response patterns when analysing tetrachoric and polychoric correlation matrices under limited sample size [18]. Secondly, CFA models could hide severe misspecification issues and result in biased parameter estimation [19,20]. Accordingly, CFA model-based reliability estimations could also be highly biased [21]. Thus, exploratory structures should be preferred in many cases [18,19]. We aim to resolve these issues by complementing these analyses with a new technique for dimensionality assessment (EGA) and the novel investigation of different exploratory factor models for the SPM-LS test.

1.2.1. Parallel Analysis

Parallel analysis is one of the main tools for dimensionality assessment [17,22,23]. Either when based on principal component or factor analysis solutions, parallel analysis has repeatedly been shown to optimally detect the true underlying unidimensionality in simulation studies [23–25]. However, parallel analysis is also fallible [18,23], with different conditions affecting each version of this procedure [17,22]. Principal component factor analysis is more reliable than the factor analysis alternative for structures with a small number of factors and binary data [17,22]. Unfortunately, it tends to wrongly suggest a single component to be retained if high factor correlations are present (as expected to occur in SPM-LS; [3]). On the other hand, factor analysis-based parallel analysis could be misleading if factors are not well defined (i.e., factor loadings < 0.40; [17]), which is indeed a plausible scenario for SPM-LS scores based on [12] depiction of APM variance partition. Additionally, either method presents difficulties in recovering the true dimensionality if samples < 500 are analysed (the size of [2] dataset; [17,26]). Finally, binary and categorical items presenting highly unbalanced categories (e.g., where the correct response represents 80–90% of the observed responses) could strongly affect parallel analysis performance [18,27,28].

1.2.2. Exploratory Graph Analysis

Exploratory Graph Analysis (EGA) is a statistical procedure that assesses latent dimensionality by exploring the unique relationships across pairs of variables (rather than the inter-item shared variance, as in common factor analysis; [29]). To do so, a sparse Gaussian Graphical Model is estimated (i.e., GGM) over the K precision matrix. K is the inverse of the inter-item variance-covariance matrix (i.e., $K = \Sigma^{-1}$; [30]) and it contains the partial correlations across pairs of observed variables. The sparse GMM is estimated by applying a penalization function (a common method is to select the GMM which minimises the extended Bayesian Information Criterion). After the GLASSO GMM is estimated, a walktrap clustering algorithm is applied to detect the optimal number of clusters in the network and to assign each item to a single dimension [21]. This algorithm, namely the combination of GLASSO GMM and walktrap clustering, has received the name of EGA. Although alternative versions of EGA exist, such as EGA with the triangulated maximally filtered graph approach (EGAtmfg), the former is preferred when high correlations between factors are expected (being the case for SPM-LS) [21].

EGA has been successfully applied to investigating the dimensionality of constructs such as personality [31], intelligence [32], and demonstrated to be as effective as parallel analysis when recovering true dimensionality under dichotomous data [17]. Nonetheless, EGA should be able to detect the number of underlying dimensions equal to or better than parallel analysis, even under suboptimal conditions (limited sample size; [17]). EGA is not presented as a substitute for techniques such as parallel analysis, but rather as a complementary tool to be studied in combination with them [17]. Accordingly, if parallel analysis results in indications of multidimensionality, researchers could benefit from exploring new techniques based on network analyses [30].

1.2.3. Exploratory Bi-factor Modelling

A review of the SPM literature has shown that two main factors models have been of interest: a unidimensional [2,4] and a multidimensional (bi-dimensional) solution [8]. Thus, it is legitimate to question to what extent specific sources of variance detected by parallel analysis or EGA could provide additional, meaningful information beyond g. In this sense, the bi-factor model should be the model to be evaluated [32,33]. The bi-factor model has been depicted as the best-suited model for assessing variance partition, to examine whether a structure is sufficiently unidimensional, and to measure the incremental value of potential specific factors [21,32,33]. When assessing estimated general factor strength, factor reliability should be compared using the omega hierarchical statistic (ω_H) [21,32]. Additionally, and to test the hypothesis of sufficient unidimensionality, the Explained Common Variance (i.e., ECV) and the Percentage of Uncontaminated Variances (PUC) should be compared altogether with ω_H for confirmatory models [34,35][1].

All model-based statistics are computed from a standardised factor analysis solution [32,36]. Therefore, it is necessary to ensure a proper estimation of the underlying bi-factor model in order to obtain unbiased reliability and ECV estimates. Given the difficulties for CFA models to recover complex structures (such as the bi-factor model) under realistic conditions (when cross-loadings are expected to occur; [19]), the bi-factor CFA models are often expected to produce biased parameter estimation [33]. In this context, exploratory alternatives such as EFA or Exploratory Structural Equation Modeling (i.e., ESEM) are becoming more and more widespread [37,38]. As these techniques offer model fit assessment while not imposing restrictions on the factor pattern matrix, they provide the modelling advantages of CFA while improving parameter estimation [18,39].

Exploratory bi-factor analysis (BEFA; Panel D, Figure 1) is a widely applied, compelling alternative to confirmatory bi-factor models [40]. The unique distinction between a BCFA and BEFA is that the latter allows the presence of cross-loadings for all specific factors [36] while maintaining the remaining characteristics (i.e., orthogonality between all factors). As each specific factor is still expected to be loaded by at least three indicators, variance partition, as well as the remaining BCFA characteristics, are present in a BEFA model [35]. However, how to approximate BEFA models is still a matter of debate. One of the most promising alternatives is via bi-factor target rotation, a technique applied in the BIFAD [10], the PEBI [41], or the SL-based iterative target rotation (SLi and SLiD algorithms; [36,38]).

In bi-factor target rotation, factor loadings to be minimised in the rotation procedure (i.e., items expected to have near-zero magnitude in the rotated loading matrix) are identified by giving them a zero value in the target matrix. As a convention, as general factor loadings are always freed (as each loading is expected to have a substantial load on this factor). The main issue then is to identify which loadings should be freed in the target rotation for the specific loadings. Conveniently, empirical cut-off points such as promin [42] or the procedure applied in SLiD algorithm [36] are able to select which loadings to be fixed based on each factor 's loadings distribution, and to prevent researchers

[1] Specific factor omega hierarchical and PUC are only computable for confirmatory solutions. Estimating such statistics in exploratory models would require researchers to decide which items or correlations are being considered by the specific factors.

from deciding on applying inappropriate fixed cut-off points (such as fixing all $\lambda < 0.20$; [36]). As an example, SLiD has been demonstrated to accurately recover bi-factor models in conditions under realistic conditions (i.e., cross-loadings or specific loadings of near-zero value), and to outperform more well-known methods such as the Schmid-Leiman orthogonalization, and the family of analytic rotations [43,44]. Promin-based algorithms (i.e., PEBI) has also been depicted as a compelling alternative and an improvement over alternative algorithms such as BIFAD [42]. Additionally, as the use of empirically defined target rotation is expected to improve parameter estimation, the estimation of general omega hierarchical, ECV and other model-based reliability estimates is also anticipated to be improved.

1.3. SPM-LS Dimensionality

SPM-LS dimensionality was evaluated by using a combination of parallel analysis, EFA and CFA results [2]. However, due to the limited sample size and the unbalanced responses patterns, parallel analysis results presented by the authors should be examined with caution. As the authors acknowledged, SPM-LS data presented some strong ceiling effects, when "10.4% of the sample had a perfect score of 12" [2] (p.114). This situation could have resulted in suboptimal performance of parallel analysis. In the results section, the authors declared that up to five factors should be retained via factor analysis parallel analysis. Additionally, and due to the large ratio of the first to second eigenvalue (5.92 to 0.97), evidence of a robust general factor was said to be found [2]. However, as factor analysis parallel analysis could be more unreliable than its principal-component alternative for the study at hand (due to limited sample size and the binary nature of the data), the results of both techniques should have been taken into consideration (e.g., when computing ratios of eigenvalues).

The authors additionally reported that no evidence of relevant specific factors was identified, as factor pattern loadings on unreported solutions including two to five factors were not in line with any theoretical expectation (i.e., "were uninterpretable"; [2], p. 112). However, the authors did not report the structures tested, or if models combining general and specific sources of variation (i.e., bi-factor) were estimated. Lastly, as global fit indexes suggested an adequate fit for the unidimensional model (i.e., even though RMSEA was as high as 0.079) and the general factor was considered as reliable ($\omega_H = 0.86$), the authors concluded that the SPM-LS scores could be considered essentially unidimensional [2] (p.112). In this investigation, this claim will be revisited by a more nuanced inspection of SPM-LS scores by applying traditional methods (exploratory and confirmatory unidimensional and bi-dimensional factor models) as well as two recently developed methods for assessing and validating multidimensional scales (EGA and bi-factor exploratory modelling).

2. Materials and Methods

2.1. Instrument and Data

The SPM-LS scores are those made publicly available by [2] for this special edition. In detail, the sample is composed of the answers of 499 undergraduate students who responded to the SPM-LS. The SPM-LS consists of the last 12 matrices the Standard Progressive Matrices [1] (i.e., those of greatest difficulty). Noteworthy, even though these items could be considered as polytomous, and essential information could be retrieved if they were treated as such [2], it is common to score them as dichotomous items: either a respondent identified the correct answer or not according to the item key provided by the authors. Accordingly, the tetrachoric correlation matrix was here studied. In this application, respondents had no time limit to complete the 12 items and were encouraged to respond to each item. Accordingly, no missing data were observed.

2.2. Statistical Analysis Plan

The following analysis will be performed to inspect the factor structure of the SPM-LS: Firstly, the dimensionality of the SPM-LS will be assessed applying both, principal component and factor analysis parallel analysis. Secondly, these results will be contrasted with those of EGA. If the SPM-LS is regarded as multidimensional, the hypothesis of essential unidimensionality will be tested by inspecting a series of unidimensional, exploratory and confirmatory bi-dimensional and bi-factor models (Figure 1). These models would be compared in terms of model fit, factor pattern results, ω_H and ECV, and PUC values (when possible). To estimate BEFA models, a bi-factor target rotation would be defined from bi-dimensional EFA solution, using the empirical cut-off point definition algorithm included in SLiD [36] and the promin cut-off estimation [42].

Most analyses were conducted in R 3.5.2. [45] in a reproducible manner using the rmarkdown [46] and the papaja [47] packages. The correlation matrix was obtained using the *cor_auto ()* function in the qgraph package [48], which provided similar results to the *tetrachoric ()* function from the psych package [49]. Principal component and factor analysis were conducted using the *fa.parallel ()* function in the psych package [49]. EGA was applied using the EGA package [50]. EFA and CFA models were computed using the lavaan package [51]. Cronbach's α and omega estimates were computed from the *reliability ()* function from the semTools package [52] following current recommendations on the field [53]. EFA models were rotated using oblique target rotation using the gradient projection algorithm included in the GPArotation package [54]. Bi-factor target was defined using the promin rotation [42] and the algorithm included in the SLiD [36]. The bi-dimensional EFA model was computed using minimum residual as the extraction method and target rotation towards the expected EGA solution. ESEM models for estimating bi-dimensional EFA and bi-factor EFA models with a free residual correlation were fitted in Mplus 7.3. Scripts for reproducing all analyses (i.e., main text, Appendices A and B results) can be found as Supplementary Data.

3. Results

3.1. Descriptive Analysis

A characteristic of the SPM-LS is that the chosen items represent the most difficult items from the SPM. However, the proportion of correct responses did not monotonically decrease as a function of item position (Figure 2), as it could be somewhat expected. The first six items (SMP1 to SMP6) had high correct proportions of correct responses ($0.76 < p_{correct} < 0.91$; where $p_{correct}$ is the observed proportion of correct answers) and were identified to present similar rates of unbalanced response patterns. On the other hand, the last three less than half of the responses collected were correct items (SPM10: $p_{correct} = 0.39$; SPM11: $p_{correct} = 0.36$ and SPM12: $p_{correct} = 0.32$). As said before, these unbalanced response patterns could lead to significant estimation errors in the tetrachoric correlation estimation.

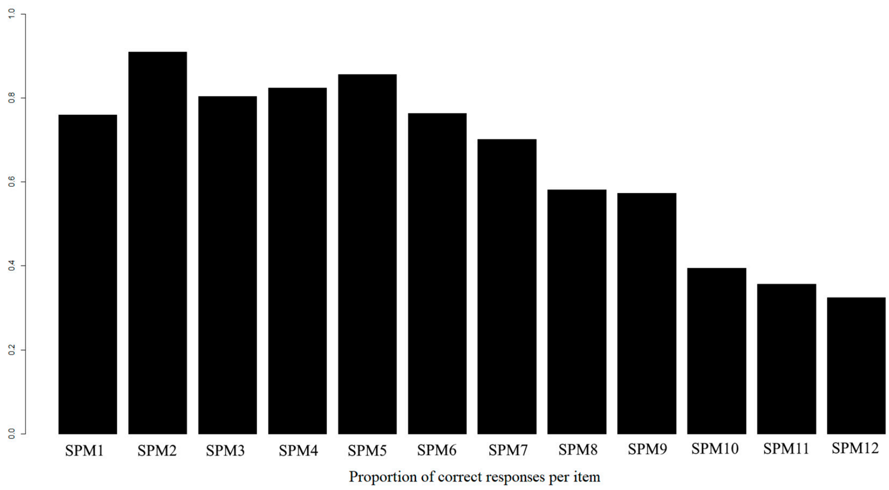

Figure 2. Proportion of correct responses as a function of item location in the SPM-LS.

A visual inspection of the tetrachoric correlation matrix (Figure 3) revealed an unusually high correlation between items ($r_{\text{SPM4}-\text{SPM15}} = 0.91$), which was substantially larger than the ensuing correlation in terms of magnitude ($r_{\text{SPM5}-\text{SPM16}} = 0.77$). In detail, 79.8% of individuals who correctly responded SPM4, also were correct for SPM5. Moreover, 11.8% of respondents who failed SPM4, also failed SPM5. Thus, there was only 8.4% of respondents who failed/gave a correct answer or gave a correct answer/failed SPM4-SPM5, respectively. A visual inspection of the tetrachoric correlation heatmap revealed two distinct blocks of inter-item correlations: The first one between items SMP1 to SPM6, and the second one between items SPM7 to SMP11. Therefore, Figure 3 is indicative of two distinct sources of multidimensionality. Due to the limited sample size, and the highly unbalanced response patterns for items such as SPM2, SPM11, and SPM12, it is noteworthy that the tetrachoric correlations between these items could be affected by significant estimation errors.

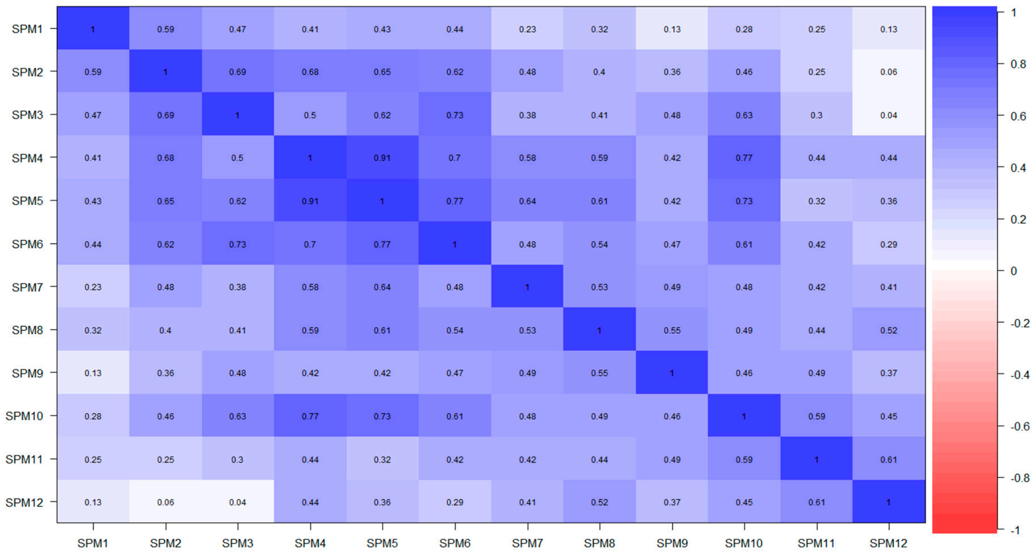

Figure 3. Heatmap of SPM-LS items tetrachoric correlation.

3.2. *Dimensionality Assessment.*

We exactly replicated the results provided by [2] when computing parallel analysis over the tetrachoric correlation matrix (using maximum likelihood)[2] (Left panel, Figure 4; also Figure 1 in [2]). The number of factors to be retained was 5, with eigenvalues of 5.92, 0.93, 0.36, 0.18, and 0.10 (simulated eigenvalues of.52, 0.21. 0.16, 0.12, 0.07). The number of components to be retained was 2, with eigenvalues as of 6.36 and 1.60 (simulated eigenvalues of 1.26 and 1.20). Noteworthy, it was observed that the authors conducted this analysis over the tetrachoric correlation matrix, obtaining the eigenvalues to be compared against those extracted by generating random normal data. However, this strategy is considered highly inadequate [18]. A better strategy when analyzing tetrachoric correlations is to obtain the random eigenvalues by resampling from the observed data. Accordingly, we repeated the analysis with this specification (Right panel, Figure 4). Factor and principal component factor analysis suggested to retain two and three factors/components, respectively: factor analysis parallel analysis showed eigenvalues of 3.43, 0.73 and 0.33 (with resampled eigenvalues of 0.54, 0.20 and 0.15) while principal components PA resulted in eigenvalues of 4.09, 1.51 for the original components (with resampled components of 1.26 and 1.19).

---

2 Using other extraction methods (i.e., ordinary least squares) led to similar conclusions regarding the underlying dimensionality, but for weighted and generalized least squares, which suggested to retain three factors and two components.

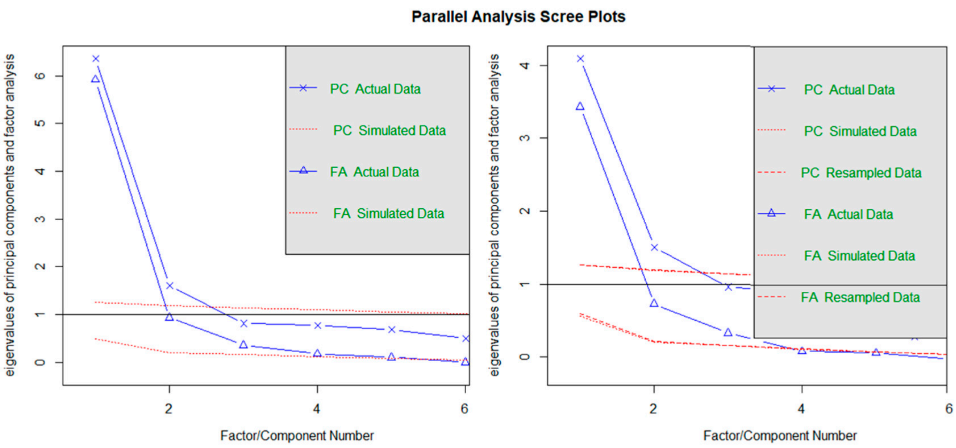

Figure 4. Parallel analysis results: (**a**) Original Principal component and parallel factor analysis with eigenvalue simulated from random normal data; (**b**) Principal component and parallel factor analysis correct eigenvalues obtained from resampling from original data.

Nevertheless, both parallel analysis techniques are suggesting the SPM-LS be multidimensional. The discrepancy between both methods (suggestions of three factors vs two components to be retained) could be due factor analysis-based parallel analysis being more affected by the limited sample size analysed. EGA agreed with principal component parallel analysis and identified two underlying dimensions (Figure 5), one composed of items one to six and the other of items seven to twelve. Moreover, EGA results confirmed that the highest observed partial correlation was observed for the pair SPM4–SPM5. This partial correlation indicates that, after controlling for all the other variables, these items were strongly conditionally dependent.

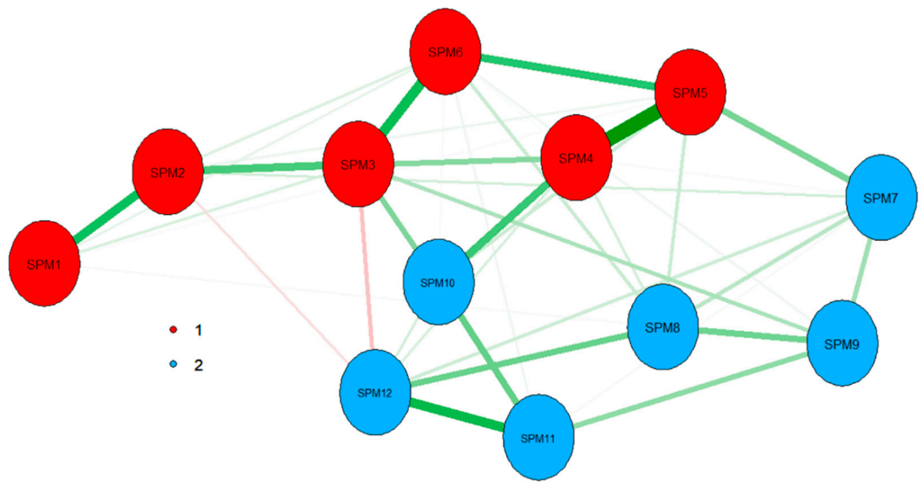

Figure 5. Exploratory Graph Analysis of SPM-LS data. Dimensions and items associated are presented in different colours. Positive partial correlations are depicted in green, with negative partial correlations presented in red. The size of the lines indicates the size of the partial correlations.

Therefore, and after inspecting the tetrachoric correlation matrix and observing the dependence between SPM4–SPM5 items, it was decided to reanalyse SPM-LS dimensionality after aggregating these items. Item parcelling (i.e., aggregating items) have been shown as a valid alternative to deal with residual item covariances [55]. Both techniques of parallel analysis agreed in this re-analysis that two factors should be retained. EGA also resulted in two factors being identified, with a similar distribution

J. Intell. **2019**, *7*, 14

than in Figure 5. Therefore, robust evidence from both, parallel analysis and EGA, supported the hypothesis of SPM-LS being bi-dimensional (either when treating the original set of items, or the reduced version combining items SPM4 and SPM5). Analysis details and results of this analysis are presented in Appendix A.

3.3. Factor Modelling

The standardised factor solutions for all estimated models are shown in Table 1. Likewise, the fit indices for all estimated models are presented in Table 2. For the sake of comparison, similar models not estimating the residual correlation between SPM4–SPM5 were also computed. Standardised factor loadings and model fit indices of these models without including this residual correlation are presented in Appendix B.

Table 1. Standardized factor loadings for all model tested.

	Unidim.	**Unidim.M.**	**BID.EFA**		**BID.CFA**		**BEFA**		**BCFA**	
Item	G	G	S1	S2	S1	S2	G	S1	G	S
SPM1	**0.47**	**0.48**	**0.59**	−0.08	**0.50**	0.00	**0.54**	−0.12	**0.50**	0.00
SPM2	**0.72**	**0.74**	**0.93**	−0.15	**0.76**	0.00	**0.82**	−0.22	**0.77**	0.00
SPM3	**0.72**	**0.73**	**0.88**	−0.09	**0.76**	0.00	**0.81**	−0.16	**0.76**	0.00
SPM4	**0.92**	**0.84**	**0.55**	**0.41**	**0.89**	0.00	**0.81**	0.25	**0.88**	0.00
SPM5	**0.94**	**0.87**	**0.67**	**0.30**	**0.91**	0.00	**0.86**	0.15	**0.91**	0.00
SPM6	**0.81**	**0.83**	**0.75**	0.17	**0.85**	0.00	**0.85**	0.05	**0.85**	0.00
SPM7	**0.66**	**0.67**	**0.30**	**0.47**	0.00	**0.71**	**0.60**	**0.33**	**0.62**	**0.30**
SPM8	**0.70**	**0.71**	0.23	**0.58**	0.00	**0.75**	**0.61**	**0.42**	**0.62**	**0.40**
SPM9	**0.61**	**0.61**	0.20	**0.50**	0.00	**0.65**	**0.53**	**0.36**	**0.52**	**0.39**
SPM10	**0.79**	**0.80**	**0.43**	**0.48**	0.00	**0.85**	**0.74**	**0.32**	**0.76**	0.27
SPM11	**0.62**	**0.63**	−0.04	**0.75**	0.00	**0.66**	**0.44**	**0.58**	**0.44**	**0.63**
SPM12	**0.53**	**0.54**	**−0.38**	**1.00**	0.00	**0.57**	0.28	**0.80**	**0.31**	**0.73**
φ	-	-	0.56		0.82		0.00		0.00	
SPM4-SPM5	-	0.69	0.70		0.56		0.70		0.57	

[1] Unidim = Unidimensional model. Unidim.M. = Unidimensional model with SPM4-SPM5 residual correlation estimated. BID.EFA = Bi-dimensional exploratory factor analysis. BID.CFA = Bi-dimensional confirmatory factor analysis. BEFA = Bi-factor exploratory factor analysis. BCFA = Bi-factor confirmatory factor analysis. All loadings over 0.30 are presented bolded. φ = Inter-factor correlation. SPM4–SPM5 = Residual covariance between SPM4-SPM5 items. G = General Factor. S1= First specific factor. S2 = Second specific factor. Factor loadings with values > 0.30 appear bolded.

Table 2. Model fit indices for all tested models.

	Np	**df**	X^2	p	**CFI**	**TLI**	**RMSEA**	**SRMR**
Unidim.	24	54	221.75	0.00	0.95	0.93	0.08 (0.07–0.09)	0.11
Unidim.M.	25	53	205.88	0.00	0.95	0.94	0.08 (0.07–0.08)	0.11
BID.EFA/BEFA.	36	42	**80.50**	**0.00**	**0.99**	**0.98**	**0.04 (0.03–0.06)**	**0.06**
BID.CFA	26	52	160.69	0.00	0.96	0.96	0.07 (0.05–0.08)	0.09
BCFA	31	47	113.72	0.00	0.98	0.97	0.05 (0.04, 0.07)	0.07

[1] Unidim = Unidimensional model. Unidim.M. = Unidimensional model with SPM4–SPM5 residual correlation estimated. BID.EFA = Bi-dimensional exploratory factor analysis. BID.CFA = Bi-dimensional confirmatory factor analysis. BEFA = Bi-factor exploratory factor analysis. BCFA = Bi-factor confirmatory factor analysis. Np = Estimated number of parameters. Df = degrees of freedom. [2] = Chi-square statistic. P = p-value associated with [2] test of fit. CFI = Comparative fit index. TLI = Tucker-Lewis index. RMSEA = Root Mean Square Error of Approximation (with 95% confidence interval in parenthesis). SRMS = Standardized Root Mean Square Residual. Best fit indices presented bolded and underlined. Model fit indices for the best fitting model appear bolded.

3.3.1. Unidimensional Model

We first replicated the original results with regards to the CFA unidimensional model [2]. We found the same model fit indices (CFI = 0.95, TLI = 0.93, RMSEA = 0.08, SRMS = 0.11). Cronbach's α = 0.92 and ω_{HG} = 0.83 also matched those reported. For this model, the high RMSEA and SRMR values suggest questionable fit. Estimating the correlation between SPM4–SPM5 resulted in improved model fit (CFI = 0.95, TLI = 0.94, RMSEA = 0.08, SRMS = 0.11). As expected, the SPM4–SPMP5 correlation was high and positive ($\psi = 0.69$). Accordingly, the remaining presented models will include the estimation of the residual correlation between both items. Additionally, this unidimensional model showed adequate reliability (Cronbach's $\alpha = 0.92$; $\omega_{HG} = 0.86$).

3.3.2. Bi-Dimensional Model

Two bi-dimensional structures were computed. Firstly, an exploratory bi-dimensional model was fitted in order to understand if EFA results supported the idea of a bi-dimensional SPM-LS structure. Secondly, such an EFA structure was tested as a confirmatory model to understand the role of potential cross-loadings present on the data. EFA model fit indexes revealed that this structure provided an excellent fit to the data (CFI = 0.99, TLI = 0.98, RMSEA = 0.04, SRMS = 0.06), improving model fit with respect to the unidimensional case. Additionally, a lower inter-factor correlation of ($\varphi \approx 0.56$) was obtained[3]. The SPM4–SPM5 correlation of this residual correlation ($\psi = 0.70$) was similar to the one observed in the unidimensional model.

The confirmatory bi-dimensional (CFI = 0.96, TLI = 0.96, RMSEA = 0.06, SRMS = 0.09) presented a better model fit than the unidimensional model, but worse than its exploratory counterpart. Fixing all cross-loadings to zero led to observe a larger factor correlation ($\varphi = 0.82$), larger SPM4–SPM5 loadings ($\lambda_{SPM4} = 0.89, \lambda_{SPM5} = 0.91$), and a diminished residual correlation between them ($\psi = 0.56$). In this case, both factors were considered as reliable if measured by Cronbach's α standards (factor 1 = 0.91, factor 2 = 0.85), and close to acceptable reliability when inspecting ω_{HS} (factor 1 = 0.75 factor 2 = 0.70). In conclusion, a bi-dimensional model (either by EFA/CFA based) improved model fit over the unidimensional structure. As indicated by the substantial inter-factor correlation observed in all models, a general factor could play a substantial role in SPM-LS structure. This hypothesis will be explored next via bi-factor modelling.

3.3.3. Bi-Factor Model

Two bi-factor models were tested: a BEFA model fitted using bi-factor target rotation and a BCFA model restricting cross-loadings to zero. Either using the algorithm included in SLiD [36] or a promin-based cut-off [42] resulted in items SPM7 to SPM12 being freed in the specific factor. Noteworthy, as rotation does not affect model fit [29], fit indices for this model were those of the exploratory bi-dimensional structure. The BEFA model (Table 1) presented three main characteristics: (a) The rotation procedure recovered orthogonal factors (even if oblique target rotation was applied), which aligns with the expectations of the bi-factor model; (b) Although the general factor was well-defined (all loadings over $\lambda_G > 0.30$), SPM11 and SPM12 presented higher loadings on the specific factor ($\lambda_{SSPM11} = 0.58, \lambda_{SSPM12} = 0.80$) than in the general factor ($\lambda_{gSPM11} = 0.44, \lambda_{GSPM12} = 0.29$); (c) the residual correlation between SPM4 and SPM5 was similar to the one observed for the unidimensional model ($\psi = 0.70$). With regards to BEFA general factor reliability, it was considered as adequate ($\omega_{HG} = 0.80$; ECV = 0.74).

The BCFA model showed the best fit indexes from all confirmatory models (Table 2; CFI = 0.98, TLI = 0.97, RMSEA = 0.05, SRMS = 0.07). Both factors were well-defined (all loadings $\lambda > 0.30$) with SPM4–SPM5 general loadings being stronger than in the BEFA model (as they were inflated due their cross-loadings being fixed to zero). SPM4–SPM5 residual correlation was similar to the one observed in the confirmatory bi-dimensional model ($\psi = 0.57$). Overall, general factor reliability was also adequate ($\omega_{HG} = 0.75$; ECV = 0.80). Additionally, the associated PUC was $(132 - 42)/132 = 0.68$. Under the presence of $PUC < 0.80$, researchers are recommended that $\omega_H > 0.70$ and $ECV > 0.60$

[3] Using alternative oblique rotations (i.e., oblimin, promax, geomin) resulted in factor structures with a similar distribution of loadings and size. Main differences were small in magnitude, and mostly affected the inter-factor correlation size.

be used as benchmarks for considering essential unidimensionality [34]. Therefore, while the BCFA provided an adequate approximation towards SPM-LS multidimensionality, the presence of a strong, reliable general factor also favours that SPM-LS scores be considered as essentially unidimensional. Lastly, the specific factor reliability (ω_{HS} = 0.31) was in the range of values commonly observed on bi-factor modelling [32,33].

4. Discussion

The SPM-LS (Standard Progressive Matrices–Last Series) has been recently proposed as an improved short version of the SPM test [2]. The SPM-LS was treated as an essentially unidimensional measure of g, with better psychometric properties than alternative tests such as the Advanced Progressive Matrices test (i.e., APM). On these grounds, [2] proceeded to fit a series of IRT models to study the benefits of studying the nominal responses in the test, acknowledging that mixed results from EFA and CFA results could suggest SPM-LS not being a strictly unidimensional measure. The authors further recommended investigators to conduct additional research on this matter. We aimed to shed light on SPM-LS dimensionality using improving the dimensionality techniques applied (comparing parallel analysis with exploratory graphic analysis results) and by providing a thoughtful exploration of unidimensional, bi-dimensional and bi-factor SPM-LS structures.

The main result of this study is that SPM-LS can be considered as essentially unidimensional measurement of intelligence if appropriately treated. Reliability and unidimensionality indices obtained from a bi-dimensional bi-factor model provided strong evidence of this conclusion. Notwithstanding the evidence of essential unidimensionality, it is also true that a non-ignorable, nuisance factor associated with the last six indicators of the SPM-LS was systematically found, either when applying parallel analysis, EGA, or factor modelling. An additional residual covariation between SPM4–SPM5 was also observed. This circumstance that should be discussed in more detail: Firstly, such a high residual correlation between both items might be due to significant estimation error in the tetrachoric matrix, altogether with the limited sample size. If so, future research employing different, larger samples should be able to identify a substantially smaller covariation between these items. Secondly, the relationship between SPM4 and SPM5 in terms of content and rules used for resolving these items should be inspected in further detail in order to decide if the information provided by both items is truly distinct or redundant.

This study evidence dimensionality assessment is a complex task which often requires convergent evidence from different sources and statistical techniques (as suggested in the case of parallel analysis and EGA; [17]). Moreover, being overconfident about model fit indices could be misleading when selecting an appropriate solution. Model fit should always be complemented with alternative indices (such as ω_H, ECV or PUC) when possible [34]. Lastly, caution should be exercised when interpreting high inter-factor correlations in confirmatory models as evidence of unidimensionality, as these correlations could be inflated if relevant cross-loadings are being omitted. As an example, the inter-factor correlation was substantially larger for the bi-dimensional confirmatory structure that for its exploratory counterpart. To avoid such situations, we recommend researchers to confront results from both exploratory and confirmatory versions of the models to be investigated. If relevant cross-loadings to be potentially fixed are identified, we agree with previous authors that exploratory models should be prioritized [19,20].

Lastly, the result of applying bi-factor modelling was clear: We found evidence of a robust and reliable g factor (which resulted in our conclusion of SPM-LS scores being essentially unidimensional by current benchmarks [34]), plus an additional nuisance factor related with the last six items. While the interpretation of this latter factor could be somewhat controversial, it cannot be associated with a speed factor as in previous applications of similar tests [7,56] (as respondents had no time limit to reply to the matrices). An alternative explication is that such a factor would be related to guessing strategy or a difficulty component. Noteworthy, the first six items were (almost uniformly) correctly responded (with a proportion of correct responses near to 0.80), with the last six items presented a decreasing proportion of right answered (as evidenced in Figure 2). Under these conditions, it is known that parallel analysis is set to fail and that exploratory factor analysis under tetrachoric correlations could result in reflecting a difficulty factor [57,58]. Alternatively, the idea of guessing strategies being a relevant aspect of SPM-LS data was strongly supported by the original authors [2], as they showed that a three-parameter IRT model (incorporating a pseudo-guessing parameter) fitted the data better than alternative models. In this sense, and as pointed out by a reviewer, statistical artefacts of similar nature could be observed when applying factor analysis to a tetrachoric correlation matrix obtained from data generated from a three-parameter IRT model. Therefore, additional research on this matter should be granted in future SPM-LS applications. Thus, evidence suggests that guessing could play a substantive role with regards to general intelligence estimation. Even though we expanded these findings by identifying that guessing could also affect dimensionality assessment, future research should focus on re-assessing SPM-LS dimensionality under the assumption of data being generated from the three-parameter nested logistic model, as it has been shown to improve the effectiveness of parallel analysis [58]. Lastly, specific item position and item difficulty effects should aim to be separately studied (as they are confounded in the current SPM-LS form). Additionally, structural models aimed to measure each specific effect should also be encouraged to be applied [14].

Overall, the consequences of the presented findings are two-folded: firstly, even though researchers could treat SPM-LS as essentially unidimensional, this does not preclude them to not use the better measurement model (i.e., the bi-factor form) in their statistical analyses, especially if included within an SEM framework. Failing to take the influence of the second factor into account could lead to inflating or deflated regression coefficient and other types of measurement error propagation [39]. As an example, in our results, the variance explained by the second factor is of 0.17. If we assume a criterion Y, measured with reliability of one and a perfect positive relationship with the nuisance factor, the expected value for the estimated correlation between our nuisance factor and Y would be estimated as 0.41 (considering the attenuation by reliability described in [59]). Even though such distorting effect represents a worst-case scenario, where expected attenuation effects are anticipated to be smaller (as either criterion reliability or true relationship between criterion or specific factor would be not perfect), they should not be disregarded as negligible [59].

An attenuation of this magnitude could impact the evaluation of SPM-LS scores criterion and incremental validity (the expected increment of the determination coefficient might range from zero to 0.17). Note that our analysis identifies a source of performance variance. The effects might be even more substantial for a group with larger variance in the secondary factor. Consequently, despite the essential unidimensionality of the measure, the consequences of taking or not this second factor into account must be weighted in future research endeavours, including additional intelligence and ability measures.

Secondly, and from a theoretical point of view, researchers should not automatically disregard such secondary factors, as they could be tied to relevant individual differences of the test-takers [15,16]. On the contrary, more research is needed for us to have a better understating of the nature of this nuisance factor, and the extent that it could represent valuable information of the examinees.

5. Conclusions

The SPM-LS has been suggested to be a valid, reliable alternative version of the Standard Progressive Matrices test, presenting superior psychometric properties to alternatives such as the Advanced Progressive Matrices test. In this research, we provided a detailed study of the essential unidimensionality claimed by the original authors by utilising applying modern dimensionality techniques and bi-factor modelling. Our results suggest that, if appropriately treated, SPM-LS scores can be considered as such. Nevertheless, an additional factor relevant to the last six items was identified. Additionally, we recommend evaluating further the presence of this factor in additional, larger sample sizes presenting more balanced responses to the SPM-LS test.

Supplementary Materials: The following are available online at http://www.mdpi.com/2079-3200/7/3/14/s1.

Author Contributions: E.G.-G.: Conceptualized the study; E.G.G., F.J.A., & L.E.G.: Established the overarching framework; E.G.-G., F.J.A., & L.E.G.: conducted the statistical analysis; E.G.-G.: Wrote the manuscript; F.J.A., & L.E.G.: Editing & approved the manuscript.

Funding: This research was by Ministry of Education, Culture and Sports (Spain) under Grant FPU15/03246 and by Ministerio de Economia y Competitividad (Spain) and Fondo Social Europeo under Grant PSI2013-44300-P and Grant PSI2017-85022-P.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

In this Appendix A, the SPM-LS dimensionality will be re-analysed by including a parcel created by aggerating SPM4-SPM5 items. This decision was taken based on the high dependence observed between items SPM4–SPM5 (i.e., tetrachoric correlation of 0.91; high partial correlation detected in EGA) Thus, we will follow the same steps performed in the primary analysis. Firstly, we reproduce the tetrachoric-polychoric correlation analysed in these analyses. As expected, most correlations between items and the combined item (i.e., SPM4-5) were like the original (Table A1).

Table A1. Tetrachoric/polychoric correlation matrix with SPM4 and SPM5 combined.

	1	2	3	4–5	6	7	8	9	10	11	12
SPM1	1										
SPM2	0.59	1									
SPM3	0.47	0.69	1								
SPM4-5	0.40	0.67	0.54	1							
SPM6	0.44	0.62	0.73	0.72	1						
SPM7	0.23	0.48	0.38	0.62	0.48	1					
SPM8	0.32	0.40	0.41	0.60	0.51	0.53	1				
SPM9	0.13	0.36	0.48	0.41	0.47	0.49	0.55	1			
SPM10	0.28	0.46	0.63	0.77	0.61	0.48	0.49	0.46	1		
SPM11	0.25	0.25	0.31	0.42	0.42	0.42	0.44	0.49	0.59	1	
SPM12	0.13	0.06	0.04	0.43	0.29	0.41	0.52	0.37	0.45	0.61	1

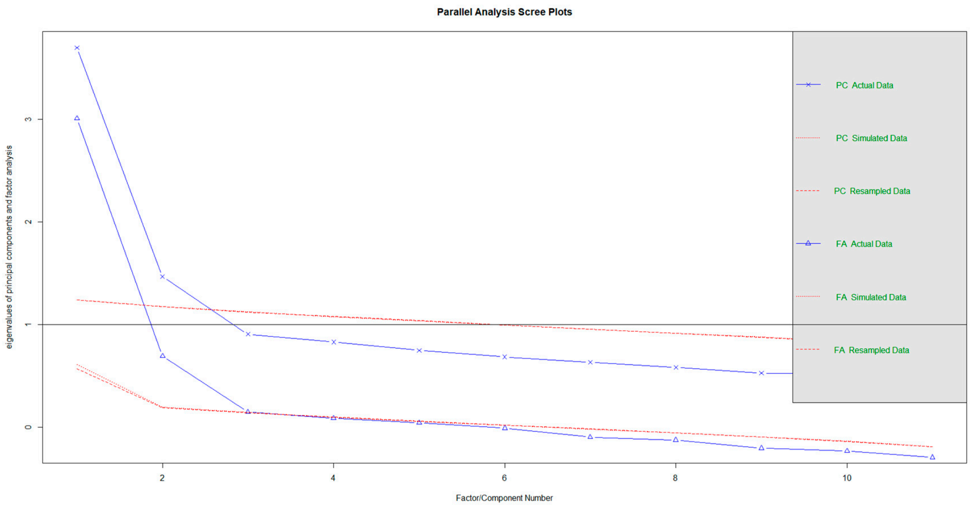

Figure A1. Principal component and parallel factor analysis with eigenvalue obtained from resampling from original data using a parcel for SPM4 and SPM5 items.

We performed principal components, and factor analysis parallel analysis with eigenvalues resampled from the original data over this correlation matrices. Both techniques agreed to indicate that the structure was bi-dimensional (Figure A2). The value of the original components was 3.70 and 1.47 (with resampled components of 1.24 and 1.17), and the value of the original factor was 3.01 and 0.69 (with resampled eigenvalues of 0.64 and 0.19).

EGA agreed with parallel analysis results and concluded that two dimensions are underlying the SPM-LS scores if SPM4 and SPM5 items were combined. Thus, there was robust evidence of the bi-dimensional nature of the data after controlling for the dependency between SPM4 and SPM5 items.

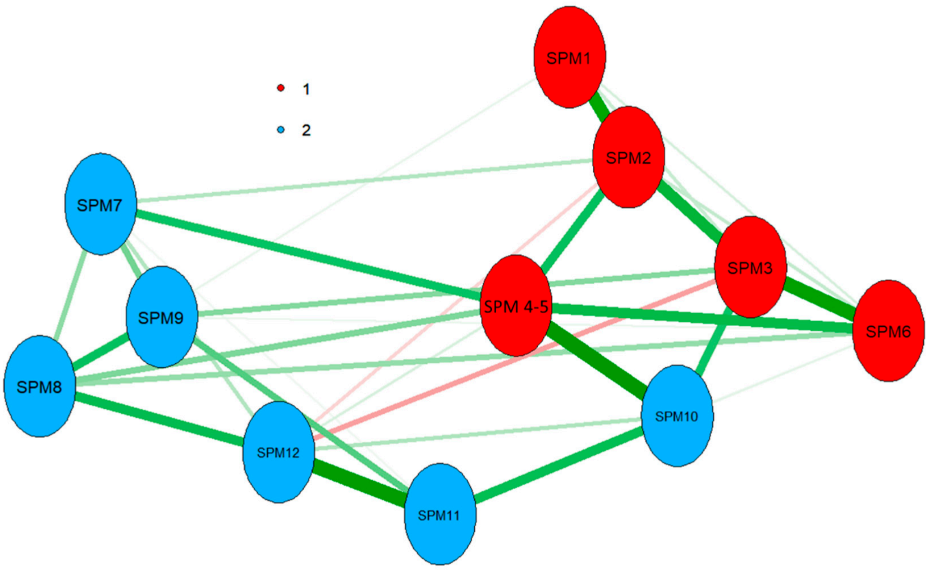

Figure A2. Exploratory Graph Analysis of SPM-LS data with SPM4 and SPM5 item combined. Dimensions and items associated are presented in different colours. Positive partial correlations are depicted in green, with negative partial correlations presented in red. The size of the lines indicates the size of the partial correlations. SPM4 = SPM4-5 item.

Table A2. Standardised factor loadings for all model tested.

	Unidim.	BID.EFA		BID.CFA		BEFA		BCFA	
Item	G	S1	S2	S1	S2	G	S1	G	S
SPM1	**0.47**	**0.58**	−0.04	**0.50**	0.00	**0.55**	−0.10	**0.51**	0.00
SPM2	**0.72**	**0.90**	−0.10	**0.76**	0.00	**0.84**	−0.17	**0.77**	0.00
SPM3	**0.74**	**0.87**	−0.04	**0.77**	0.00	**0.84**	−0.13	**0.79**	0.00
SPM4-5	**0.85**	**0.55**	**0.43**	**0.90**	0.00	**0.82**	0.28	**0.90**	0.00
SPM6	**0.82**	**0.70**	0.22	**0.86**	0.00	**0.84**	0.10	**0.85**	0.00
SPM7	**0.67**	0.26	**0.51**	0.00	**0.70**	**0.57**	**0.36**	**0.60**	**0.31**
SPM8	**0.71**	0.20	**0.60**	0.00	**0.74**	**0.58**	**0.45**	**0.61**	**0.41**
SPM9	**0.62**	0.19	**0.52**	0.00	**0.65**	**0.52**	**0.38**	**0.52**	**0.40**
SPM10	**0.80**	**0.40**	**0.51**	0.00	**0.84**	**0.72**	**0.35**	**0.75**	0.29
SPM11	**0.64**	−0.04	**0.75**	0.00	**0.67**	**0.43**	**0.59**	**0.45**	**0.61**
SPM12	**0.54**	**−0.39**	**1.00**	0.00	**0.57**	0.23	**0.82**	0.29	**0.76**
φ	**-**	**0.54**		**0.81**		0.00		0.00	

[1] Unidim = Unidimensional model. Unidim.M. = Unidimensional model with SPM4-SPM5 residual correlation estimated. BID.EFA = Bi-dimensional exploratory factor analysis. BID.CFA = Bi-dimensional confirmatory factor analysis. BEFA = Bi-factor exploratory factor analysis. BCFA = Bi-factor confirmatory factor analysis. All loadings over 0.30 are presented bolded. φ = Inter-factor correlation. SPM4-SPM5 = Residual covariance between SPM4-SPM5 items. G = General Factor. S1= First specific factor. S2 = Second specific factor. Factor loadings with values > 0.30 appear bolded.

Table A3. Model fit indices for all tested models.

	Np	df	X^2	*p*	CFI	TLI	RMSEA	SRMR
Unidim.	23	44	192.04	0.00	0.93	0.91	0.08 (0.07–0.09)	0.11
BID.EFA/BEFA.	33	34	68.45	0.00	**0.98**	**0.97**	**0.05 (0.03–0.06)**	**0.05**
BID.CFA	24	43	145.71	0.00	0.95	0.94	0.07 (0.06–0.08)	0.09
BCFA	29	38	110.79	0.00	0.97	0.96	0.06 (0.04–0.07)	0.07

[1] Unidim = Unidimensional model. Unidim.M. = Unidimensional model with SPM4-SPM5 residual correlation estimated. BID.EFA = Bi-dimensional exploratory factor analysis. BID.CFA = Bi-dimensional confirmatory factor analysis. BEFA = Bi-factor exploratory factor analysis. BCFA = Bi-factor confirmatory factor analysis. Np = Estimated number of parameters. Df = degrees of freedom. [2] = Chi-square statistic. P = *p*-value associated with [2] test of fit. CFI = Comparative fit index. TLI= Tucker-Lewis index. RMSEA = Root Mean Square Error of Approximation (with 95% confidence interval in parenthesis). SRMS = Standardized Root Mean Square Residual. Best fit indices presented bolded and underlined. Model fit indices for the best fitting model appear bolded.

Lastly, and in the case to be of interest, standardised factor loadings and model fit indices are provided. Noteworthy, results were similar to other models presented in this article but provided a sustainably worse fit to the data. In the exploratory models, SPM4-5 showed lower factor loadings in the S1 (model BID.EFA) or G (model BEFA), and higher cross-loadings on the alternative factors. In the confirmatory models, SPM4-5 loadings were also closer to 0.90 than in the main text results. Overall, resulting structures were mostly similar to those analysed in the result section of the article.

Appendix B

In Appendix B, standardised factor loadings (Table A4) and model fit indices (Table A5) are provided for models without the residual correlation SPM4-SPM5.

Table A4. Standardised factor loadings for all model tested.

	Unidim.	BID.EFA		BID.CFA		BEFA		BCFA	
Item	G	S1	S2	S1	S2	G	S1	G	S
SPM1	**0.47**	**0.59**	−0.09	**0.50**	0.00	**0.53**	−0.13	**0.50**	0.00
SPM2	**0.72**	**0.93**	−0.18	**0.75**	0.00	**0.81**	−0.23	**0.76**	0.00
SPM3	**0.72**	**0.87**	−0.10	**0.75**	0.00	**0.80**	−0.16	**0.76**	0.00
SPM4	**0.92**	**0.65**	**0.38**	**0.94**	0.00	**0.90**	0.22	**0.93**	0.00
SPM5	**0.94**	**0.74**	**0.30**	**0.95**	0.00	**0.94**	0.16	**0.96**	0.00
SPM6	**0.81**	**0.73**	0.15	**0.84**	0.00	**0.83**	0.04	**0.84**	0.00
SPM7	**0.66**	**0.30**	**0.47**	0.00	**0.71**	**0.59**	**0.33**	**0.61**	**0.31**
SPM8	**0.70**	0.23	**0.57**	0.00	**0.75**	**0.60**	**0.42**	**0.61**	**0.41**
SPM9	**0.60**	0.20	**0.50**	0.00	**0.64**	**0.52**	**0.36**	**0.51**	**0.40**
SPM10	**0.79**	**0.43**	**0.47**	0.00	**0.84**	**0.73**	**0.31**	**0.75**	**0.30**
SPM11	**0.62**	-0.05	**0.75**	0.00	**0.66**	**0.44**	**0.58**	**0.44**	**0.64**
SPM12	**0.53**	**−0.36**	**0.99**	0.00	**0.57**	0.28	**0.80**	**0.31**	**0.72**
φ	**-**	**0.57**		**0.80**		0.00		0.00	

[1] Unidim = Unidimensional model. Unidim.M. = Unidimensional model with SPM4-SPM5 residual correlation estimated. BID.EFA = Bi-dimensional exploratory factor analysis. BID.CFA = Bi-dimensional confirmatory factor analysis. BEFA = Bi-factor exploratory factor analysis. BCFA = Bi-factor confirmatory factor analysis. All loadings over 0.30 are presented bolded. Phi = Inter-factor correlation. SPM4-SPM5 = Residual covariance between SPM4-SPM5 items. G = General Factor. S1= First specific factor. S2 = Second specific factor. Factor loadings with values > 0.30 appear bolded.

Table A5. Model fit indices for all tested models.

	Np	df	X^2	*p*	CFI	TLI	RMSEA	SRMR
Unidim.	24	54	221.75	0.00	0.94	0.93	0.08 (0.08–0.09)	0.11
BID.EFA/BEFA.	35	43	97.21	0.00	**0.98**	**0.97**	**0.05 (0.04–0.06)**	**0.06**
BID.CFA	25	53	163.39	0.00	0.96	0.96	0.07 (0.05–0.07)	0.09
BCFA	30	48	117.65	0.00	0.98	0.97	0.05 (0.04–0.07)	0.07

[1] Unidim = Unidimensional model. Unidim.M. = Unidimensional model with SPM4-SPM5 residual correlation estimated. BID.EFA = Bi-dimensional exploratory factor analysis. BID.CFA = Bi-dimensional confirmatory factor analysis. BEFA = Bi-factor exploratory factor analysis. BCFA = Bi-factor confirmatory factor analysis. Np = Estimated number of parameters. Df = degrees of freedom. ^{2} = Chi-square statistic. P = *p*-value associated with ^{2} test of fit. CFI = Comparative fit index. TLI= Tucker-Lewis index. RMSEA = Root Mean Square Error of Approximation (with 95% confidence interval in parenthesis). SRMS = Standardized Root Mean Square Residual. Best fit indices presented bolded and underlined. Model fit indices for the best fitting model appear bolded.

References

1. Raven, J.C. Standardization of Progressive Matrices. *Br. J. Med. Psychol.* **1941**, *19*, 137–150. [CrossRef]
2. Myszkowski, N.; Storme, M. A snapshot of g? Binary and polytomous item-response theory investigations of the last series of the Standard Progressive Matrices (SPM-LS). *Intelligence* **2018**, *68*, 109–116. [CrossRef]
3. Abad, F.J.; Colom, R.; Rebollo, I.; Escorial, S. Sex differential item functioning in the Raven's Advanced Progressive Matrices: Evidence for bias. *Pers. Indiv. Differ.* **2004**, *36*, 1459–1470. [CrossRef]
4. Lucio, P.S.; Cogo-Moreira, H.; Puglisi, M.; Polanczyk, G.V.; Little, T.D. Psychometric Investigation of the Raven's Colored Progressive Matrices Test in a Sample of Preschool Children. *Assessment* **2017**, 1–11. [CrossRef] [PubMed]
5. Walsch, N.A.; Nettelbeck, S.A.J.; Nicholas, R.B. Dimensionality of the Raven's Advanced Progressive Matrices: Sex Differences and Visuospatial Ability. *Pers. Individ. Differ.* **2016**, *100*, 157–166. [CrossRef]
6. Lohmann, D.F.; Lakin, J.M. Intelligence and Reasoning. In *The Cambridge Handbook of Intelligence*; Sternberg, R.J., Kaufman, S.B., Eds.; Cambridge University Press, Ltd.: New York, NY, USA, 2011; pp. 419–441.
7. Dillon, R.F.; Pohlmann, J.T.; Lohman, D. A Factor Analysis of Raven's Advanced Progressive Matrices Freed from Difficulty Factors. *Educ. Psychol. Meas.* **1981**, *41*, 1295–1302. [CrossRef]
8. Bors, D.A.; Stokes, T.L. Raven's Advanced Progressive Matrices: Norms for First-Year University Students and the Development of a Short Form. *Educ. Psychol. Meas.* **1998**, *58*, 392–398. [CrossRef]
9. Holzinger, K.; Swineford, F. The Bi-factor Method. *Psychometrika* **1937**, *2*, 41–54. [CrossRef]
10. Waller, N.G. Direct Schmid-Leiman Transformations and Rank-Deficient Loadings Matrices. *Psychometrika* **2017**, *83*, 858–887. [CrossRef] [PubMed]
11. Johnson, W.; Bouchard, T.J. The MISTRA data: Forty-two mental ability tests in three batteries. *Intelligence* **2011**, *39*, 82–88. [CrossRef]
12. Gignac, G.E. Raven's is not a pure measure of general intelligence: Implications for g factor theory and the brief measurement of g. *Intelligence* **2015**, *52*, 71–79. [CrossRef]
13. Muniz, M.; Gomez, C.; Pasian, S. Factor Structure of Raven's Coloured Progressive Matrices. *Psico-USF* **2016**, *21*, 259–272. [CrossRef]
14. Zeller, F.; Reiß, S.; Schweizer, K. Is the Item-Position Effect in Achievement Measures Induced by Increasing Item Difficulty. *Struct. Equ. Model.* **2019**, *24*, 745–754. [CrossRef]
15. Ren, X.; Wang, T.; Sun, S.; Deng, M.; Scheizer, K. Speeded testing in the assessment of intelligence gives rise to speed factor. *Intelligence* **2018**, *66*, 64–71. [CrossRef]
16. Sun, S.; Scheizer, K.; Ren, X. Item-Position Effect in Raven's Matrices: A Developmental Perspective. *J. Cogn. Dev.* **2019**, 1–10. [CrossRef]
17. Golino, H.F.; Shi, D.; Garrido, L.E.; Christensen, A.; Nieto, M.D.; Sadana, P.; Thiyagarajan, J.A. Investigating the performance of exploratory graph analysis and traditional techniques to identify the number of latent factors: A simulation and tutorial. *Psychol. Methods* **2019**, 1–20. [CrossRef]
18. Lubbe, D. Parallel Analysis with Categorical Variables: Impact of Category Probability Proportions on Dimensionality Assessment Accuracy. *Psychol. Methods* **2018**, *24*, 339–351. [CrossRef] [PubMed]

19. Marsh, H.; Morin, A.; Parker, P.; Kaur, G. Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annu. Rev. Clin. Psychol.* **2014**, *10*, 85–110. [CrossRef] [PubMed]
20. Marsh, H.; Muthen, B.; Asparouhov, T.; Lüdke, O.; Robitzsch, A.; Morin, A.; Trautwein, U. Exploratory structural equation modelling, integrating CFA and EFA: Application to student's evaluations of university teaching. *Struct. Equ. Model.* **2009**, *16*, 439–476. [CrossRef]
21. Revelle, W.; Wilt, J. The general factor of personality: A general critique. *J. Res. Pers.* **2013**, *47*, 493–504. [CrossRef] [PubMed]
22. Timmerman, M.E.; Lorenzo-Seva, U. Dimensionality Assessment of Ordered Polytomous Items with Parallel Analysis. *Psychol. Methods* **2016**, *16*, 209–222. [CrossRef] [PubMed]
23. Garrido, L.E.; Abad, F.J.; Ponsoda, V. Are fit indices really fit to estimate the number of factors with categorical variables? Some cautionary findings via Monte Carlo simulation. *Psychol. Methods* **2016**, *21*, 93–111. [CrossRef] [PubMed]
24. Garrido, L.E.; Abad, F.J.; Ponsoda, V. A new look at Horn's parallel analysis with ordinal variables. *Psychol. Methods* **2013**, *18*, 454–474. [CrossRef] [PubMed]
25. Raiche, G.; Walls, T.; Magis, D.; Riopel, M.; Blais, J.G. Non-graphical Solutions for Cattell's Scree Test. *Methodology* **2013**, *9*, 23–29. [CrossRef]
26. Crawford, A.V.; Green, S.B.; Levy, R.; Lo, W.J.; Scott, L.; Svetina, D.; Thompson, M.S. Evaluation of parallel analysis methods for determining the number of factors. *Educ. Psychol. Meas.* **2010**, *70*, 885–901. [CrossRef]
27. Parry, D.H.; McArdle, J.J. An Applied Comparison of Methods for Least-Squares Factor Analysis of Dichotomous Variables. *Appl. Psychol. Meas.* **1991**, *15*, 35–46. [CrossRef]
28. Weng, L.J.; Cheng, C.P. Parallel Analysis with Unidimensional Binary Data. *Educ. Psychol. Meas.* **2005**, *65*, 697–716. [CrossRef]
29. Mulaik, S. *Foundations of Factor Analysis*, 2nd ed.; Chapman & Hall/CRC: Boca Raton, FL, USA, 2010.
30. Golino, H.F.; Epskamp, S. Exploratory Graph Analysis: A New Approach for Estimating the Number of Dimensions in Psychological Research. *PLoS ONE* **2017**, *12*, e0174035. [CrossRef] [PubMed]
31. Golino, H.F.; Demetriou, A. Estimating the Dimensionality of Intelligence like Data using Exploratory Graph Analysis. *Intelligence* **2017**, *62*, 54–57. [CrossRef]
32. Rodriguez, A.; Reise, S.P.; Haviland, M.G. Applying Bifactor Statistical Indices in the Evaluation of Psychological Measures. *J. Pers. Assess.* **2016**, *98*, 223–237. [CrossRef]
33. Rodriguez, A.; Reise, S.P.; Haviland, M.G. Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychol. Methods* **2016**, *21*, 137–150. [CrossRef] [PubMed]
34. Reise, S.P.; Scheines, R.; Widaman, K.; Haviland, M. Multidimensionality and Structural Coefficient Bias in Structural Equation Modelling: A Bifactor Perspective. *Educ. Psychol. Meas.* **2013**, *73*, 5–26. [CrossRef]
35. Reise, S.P.; Bonifay, W.; Haviland, M.G. Bifactor modelling and the evaluation of scale scores. In *The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale and Test Development*, 1st ed.; Irwing, P., Booth, T., Hughes, D.J., Eds.; John Wiley & Sons Ltd.: Hoboken, NJ, USA, 2018; Volume 2, pp. 675–707.
36. Garcia-Garzon, E.; Abad, F.J.; Garrido, L.E. Improving Bi-factor Exploratory Modelling: Empirical Target Rotation Based on Loading Differences. *Methodology* **2019**, *15*, 45–55. [CrossRef]
37. Mai, Y.; Zhang, Z.; Wen, Z. Comparing Exploratory Structural Equation Modeling and Existing Approaches for Multiple Regression with Latent Variables. *Struct. Equ. Model.* **2018**, *25*, 737–749. [CrossRef]
38. Abad, F.J.; Garcia-Garzon, E.; Garrido, L.E.; Barrada, J.R. Iteration of Partially Specified Target Matrices: Application to the Bi-Factor Case. *Multivar. Behav. Res.* **2017**, *52*, 416–429. [CrossRef] [PubMed]
39. Asparouhov, T.; Muthen, B. Exploratory Structural Equation Modeling. *Struct. Equ. Model.* **2009**, *16*, 397–438. [CrossRef]
40. Mansolf, M.; Reise, S.P. Exploratory Bifactor Analysis: The Schmid-Leiman Orthogonalization and Jennrich-Bentler Analytic Rotations. *Multivar. Behav. Res.* **2016**, *51*, 695–717. [CrossRef]
41. Lorenzo-Seva, U.; Ferrando, P.J. A General Approach for Fitting Pure Exploratory Bifactor Models. *Multivar. Behav. Res.* **2019**, *54*, 15–30. [CrossRef] [PubMed]
42. Lorenzo-Seva, U. Promin: A Method for Oblique Factor Rotation. *Multivar. Behav. Res.* **1999**, *34*, 347–365. [CrossRef]
43. Jennrich, R.I.; Bentler, P. Exploratory Bi-factor Analysis. *Psychometrika* **2011**, *76*, 537–549. [CrossRef]

44. Jennrich, R.I.; Bentler, P. Exploratory Bi-factor Analysis: The Oblique Case. *Psychometrika* **2012**, *77*, 442–454. [CrossRef] [PubMed]
45. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2019; Available online: https://www.R-project.org/ (accessed on 20 June 2019).
46. Allaire, J.J.; Xie, Y.; McPherson, J.; Luraschi, J.; Ushey, K.; Atkins, A.; Wickham, H.; Cheng, J.; Chang; Iannone, R. Rmarkdown: Dynamic Documents for R. R Package Version 1.12. 2019. Available online: https://rmarkdown.rstudio.com (accessed on 20 June 2019).
47. Aust, F.; Barth, M. Papaja: Prepare Reproducible APA Journal Articles with R Markdown. R Package Version 0.1.0.9842. 2018. Available online: https://github.com/crsh/papaja (accessed on 20 June 2019).
48. Epskamp, S.; Cramer, A.O.J.; Waldorp, L.J.; Schmittmann, V.D.; Borsboom, D. Qgraph: Network Visualizations of Relationships in Psychometric Data. *J. Stat. Softw.* **2012**, *48*, 1–19. Available online: http://jstatsoft.org&v48/i04/ (accessed on 20 June 2019).
49. Revelle, W. *Psych: Procedures for Personality and Psychological Research*; Northwestern University: Evanston, IL, USA, 2018; Available online: https://www.scholars.northwestern.edu/en/publications/psych-procedures-for-personality-and-psychological-research (accessed on 20 June 2019).
50. Golino, H. EGA: Exploratory Graph Analysis: Estimating the Number of Dimensions in Psychological Data. 2019. Available online: http://github.com/hfgolino/EGA (accessed on 20 June 2019).
51. Rosseel, Y. lavaan: An R Package for Structural Equation Modeling. *J. Stat. Softw.* **2012**, *48*, 1–36. Available online: http://www.jstatsoft.org/v48/i02 (accessed on 20 June 2019). [CrossRef]
52. Jorgensen, T.D.; Pornprasertmanit, S.; Schoemann, A.M.; Rosseel, Y. SemTools: Useful Tools for Structural Equation Modeling. R Package Version 0.5-1. 2018. Available online: https://CRAN.R-project.org/package=semTools (accessed on 20 June 2019).
53. Viladrich, C.; Angulo-Brunet, A.; Doval, E. A Journey Around Alpha and Omega to Estimate Internal Consistency Reliability. *Ann. Psychol.* **2017**, *33*, 755–782. [CrossRef]
54. Bernaards, C.A.; Jennrich, R.I. Gradient Projection Algorithms and Software for Arbitrary Rotation Criteria in Factor Analysis. *Educ. Psychol. Meas.* **2005**, *65*, 676–696. [CrossRef]
55. Little, T.D.; Rhemtulla, M.; Gibson, K.; Schoemann, A.M. Why the Items versus Parcels Controversy Needn't Be one. *Psychol. Methods* **2013**, *18*, 285–300. [CrossRef] [PubMed]
56. Estrada, E.; Román, F.J.; Abad, F.J.; Colom, R. Separating power and speed components of standardized intelligence measures. *Intelligence* **2017**, *61*, 159–168. [CrossRef]
57. Carroll, J.B. The effect of difficulty and chance success on correlations between items or between tests. *Psychometrika* **1945**, *10*, 1–19. [CrossRef]
58. DeMars, C.E. Revised Parallel Analysis with Nonnormal Ability and a Guessing Parameter. *Educ. Psychol. Meas.* **2019**, *79*, 151–169. [CrossRef] [PubMed]
59. Abad, F.J.; Sorrel, M.A.; Garcia, L.F.; Aluja, A. Modeling General, Specific, and Method Variance in Personality Measures: Results for ZKA-PQ and NEO-PI-R. *Assessment* **2018**, *25*, 959–977. [CrossRef]

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Chapter 5

On Omega Hierarchical Estimation

This study is to be published in the journal *Multivariate Behavioral Research*. The following pages include this publication. This article is currently in press.

MULTIVARIATE BEHAVIORAL RESEARCH
https://doi.org/10.1080/00273171.2020.1736977

Routledge
Taylor & Francis Group

Check for updates

On Omega Hierarchical Estimation: A Comparison of Exploratory Bi-Factor Analysis Algorithms

Eduardo Garcia-Garzon[a,b], Francisco J. Abad[a], and Luis E. Garrido[c]

[a]Universidad Autónoma de Madrid; [b]Universidad Camilo José Cela; [c]Pontificia Universidad Católica Madre y Maestra

ABSTRACT

As general factor modeling continues to grow in popularity, researchers have become interested in assessing how reliable general factor scores are. Even though omega hierarchical estimation has been suggested as a useful tool in this context, little is known about how to approximate it using modern bi-factor exploratory factor analysis methods. This study is the first to compare how omega hierarchical estimates were recovered by six alternative algorithms: Bi-quartimin, bi-geomin, Schmid-Leiman (SL), empirical iterative empirical target rotation based on an initial SL solution (SLiD), direct SL (DSL), and direct bi-factor (DBF). The algorithms were tested in three Monte-Carlo simulations including bi-factor and second-order structures and presenting complexities such as cross-loadings or pure indicators of the general factor and structures without a general factor. Results showed that SLiD provided the best approximation to omega hierarchical under most conditions. Overall, neither SL, bi-quartimin, nor bi-geomin produced an overall satisfactory recovery of omega hierarchical. Lastly, the performance of DSL and DBF depended upon the average discrepancy between the loadings of the general and the group factors. The re-analysis of eight classical datasets further illustrated how algorithm selection could influence judgments regarding omega hierarchical.

KEYWORDS
Bi-factor; reliability; omega; omega hierarchical; Schmid-Leiman

Introduction

The presence of general factors has been widely discussed in areas such as intelligence (Mansolf & Reise, 2016), personality (Revelle & Wilt, 2013), and psychopathology (Caspi & Moffitt, 2018). Due to its theoretical appeal, researchers in these areas have been interested in how to properly evaluate the presence of these general factors (Rodriguez et al., 2016a, 2016b). Accordingly, researchers have turned their attention to how to adequately estimate the reliability of such structures (Revelle & Condon, 2018, 2019; Revelle & Wilt, 2013).

As classical reliability indices such as Cronbach's alpha rely on untenable assumptions in most general factor modeling applications (Zinbarg et al., 2006), model-based reliability has gained in popularity. A statistic receiving considerable attention in the general factor modeling literature is the omega hierarchical coefficient (McDonald, 1999; Revelle & Zinbarg, 2009). Omega hierarchical, which was originally proposed by McDonald (1999), represents the ratio of variance accounted for by a single general factor to the test variance. Omega hierarchical has also been recommended as a tool to evaluate the presence of general factors (Revelle & Wilt, 2013) and the psychometric properties of summed or average scores (Rodriguez et al., 2016b). As previously noted, omega hierarchical can only be computed after a factor model including a general factor has been estimated. A natural option in this context is the bi-factor model (Holzinger & Swineford, 1937), as it provides a direct estimation of the contributions of general and group factors underlying the data (Rodriguez et al., 2016a, 2016b). As criticisms of bi-factor confirmatory models have gained traction (Morin et al., 2016), recent interest in bi-factor exploratory factor analysis (BEFA) has arisen. A common practice is to approximate BEFA models by employing the Schmid-Leiman transformation (SL; Schmid & Leiman, 1957; see also Revelle & Zinbarg, 2009; Rodriguez et al., 2016a, 2016b; Zinbarg & Alden, 2015). However, SL has been shown to underperform when recovering true bi-factor parameters under certain conditions (e.g., when pure indicators are present; Abad et al., 2017; Garcia-Garzon et al., 2019). Alternatively, two distinct families of algorithms have recently been

CONTACT Eduardo Garcia-Garzon eduardo.garciag@uam.es Departamento de Psicología Social y Metodología, Facultad de Psicología, Universidad Autónoma de Madrid. Ciudad Universitaria de Cantoblanco, 28049 Madrid, Spain.

Supplemental material for this article can be accessed at https://doi.org/10.1080/00273171.2020.1736977.

© 2020 Taylor & Francis Group, LLC

proposed to conduct BEFA. On the one hand, there are BEFA methods based on the adaptation of existent rotation criteria to the specific bi-factor case, bi-quartimin (Jennrich & Bentler, 2011) and bi-geomin (Jennrich & Bentler, 2012), and on the other, BEFA algorithms applying a bi-factor target rotation: Iterative empirical target rotation based on an initial SL solution (i.e., SLiD; Garcia-Garzon et al., 2019) and the Direct Schmid-Leiman (i.e., DSL) and Direct Bi-Factor (i.e., DBF) algorithms (Waller, 2018). While most of these algorithms have been compared with regard to their precision in recovering factor model parameters (Abad et al., 2017; Garcia-Garzon et al., 2019; Giordano & Waller, 2019), the usefulness of these methods in the context of model-based reliability estimates remains unexplored.

This study aims to fill this gap in knowledge by comparing the performance of the BEFA algorithms mentioned above (SL, SLiD, bi-geomin, bi-quartimin, DSL, and DBF) in the estimation of the omega hierarchical in bi-factor and second-order structures via three Monte Carlo simulations and the re-analysis of eight classical datasets. The rest of the article will be organized as follows: First, omega hierarchical and its relationship with bi-factor modeling will be introduced. Second, a brief account of the rationale behind the bi-geomin, bi-quartimin, SLiD, and DSL-DBF algorithms will be presented. Third, the results of a Monte Carlo simulation study of bi-factor and second-order structures, as well as the empirical datasets, will be displayed. Lastly, we will discuss the merits and drawbacks of each method.

For the remainder of the article, the following definitions will apply: A factor pattern matrix is said to follow a bi-factor model if and only if a general factor (i.e., a factor directly influencing $i = 1, \ldots, p$ indicators in a set of items) along with several potential group factors exists (i.e., with $j = 2, \ldots, k$ factors directly influencing $i = 1, \ldots, p_j$ subgroups of the p indicators), where the latter explain variance that is residual to the former and all factors and items' unique variances are orthogonal (Holzinger & Swineford, 1937). Additionally, general conditions for factor model identification shall hold, and the rank of such pattern matrices shall not be less than the total number of factors involved.

A bi-factor model where each item is only influenced by a general and a single group factor (resembling the concept of simple structure) is said to follow an independent cluster structure (IC; McDonald, 1999). A bi-factor model composed of items that are loaded by a general and two or more group factors (i.e., presenting cross-loadings) is said to conform to an independent cluster basis structure (ICB). If an IC structure presents one or more items that are not loaded by any group factor (i.e., pure indicators of the general factor), it is labeled an independent cluster pure structure (ICP). Factor pattern matrices combining items presenting cross-loadings and items that are pure indicators within the same group factor are denominated independent cluster basis pure structures (ICBP).

General factor model-based reliability

Recent controversies over the existence of general factors have arisen in areas such as intelligence (Mansolf & Reise, 2016), psychopathology (Caspi & Moffitt, 2018), and personality (Revelle & Wilt, 2013). As researchers' interest in general factor modeling has steadily grown, methodologists have questioned current practices used to determine whether general factors do in fact underlie a given set of items (Reise et al., 2018; Rodriguez et al., 2016a). Therefore, a common situation arises where, after fitting a general factor model to the data, researchers are concerned with assessing the extent to which a general factor accounts for total test variance (McNeish, 2018; Yang & Green, 2015).

It is a well-established fact that traditional reliability estimators such as Cronbach's alpha yield biased reliability estimates unless a set of unrealistic, stringent assumptions are met (Cronbach, 1951; Zinbarg et al., 2006). Consequently, total score reliability is currently approached using model-based reliability estimators such as omega (McNeish, 2018; Rodriguez et al., 2016a). However, in general factor modeling, a statistic receiving great attention has been the omega hierarchical, which represents the ratio of variance accounted for by a single general factor to test variance. Omega hierarchical estimation requires researchers to fit a multidimensional factor model including a general factor (Reise et al., 2018). A common choice for this factor model is the bi-factor model (Holzinger & Swineford, 1937). Compared with other alternatives, the bi-factor model provides a straightforward computation of the direct contributions of general and group factors as sources of variance (Rodriguez et al., 2016b). Following McDonald (1999), a test's variance can be decomposed into the sum of the general factor true variance, the true joint variance due to all group factors, and the error variance:

$$\sigma_X^2 = \left(\sum\nolimits_{i=1}^{p} \lambda_{ig}\right)^2 + \sum\nolimits_{j=2}^{k} \left(\sum\nolimits_{i=1}^{p} \lambda_{ij}\right)^2 + \sum\nolimits_{i=1}^{p} \theta_i \tag{1}$$

where latent factors are assumed to be standardized, σ_X^2 is the test variance, λ_{ig} is the loading for the i-th indicator on the first, general factor, λ_{ij} is the loading for the i-th indicator included in the domain on the j-th group factor, and θ_i is the i-th item uniqueness. According to Bentler (2009), the model-based implied variance ($\hat{\sigma}_X^2$) could be a more efficient estimator of the population test variance than the observed sample counterpart (V_X). Thus, this approach will hereafter be followed. Noteworthy, other authors favor the use of the original definition of omega and opt to apply the observed test variance instead (Revelle & Condon, 2019). Either way, this decision could be inconsequential if the factor model fits the data adequately. When the proportion of interest is the ratio between the true variance uniquely attributed to a general factor and the total test implied variance, it is computed as:

$$\omega_H = \frac{\left(\sum_{i=1}^{p} \lambda_{ig}\right)^2}{\left(\sum_{i=1}^{p} \lambda_{ig}\right)^2 + \sum_{j=2}^{k}\left(\sum_{i=1}^{p} \lambda_{ij}\right)^2 + \left(\sum_{i=1}^{p} \theta_i\right)} \qquad (2)$$

which is termed omega hierarchical (ω_H; Rodriguez et al., 2016b). Even though some cutoffs for omega hierarchical have been proposed, such as considering that omega values over .80 indicate that "total scores could be essentially unidimensional, in the sense that the vast majority of reliable variance is attributed to a single common source [the general factor]" (Rodriguez et al., 2016b, p. 225), researchers should be aware that the appropriateness of this cutoff is still a matter of debate in the literature. Thus, its use will be avoided in this article.

Omega hierarchical has been widely praised as a useful statistic for evaluating general factor importance (Revelle & Wilt, 2013; Zinbarg & Alden, 2015), among other uses. Accordingly, it constitutes a central statistic in general factor modeling and is extensively reported in the literature (Rodriguez et al., 2016b). Even though the properties of the omega hierarchical have been critically discussed in the literature (Hancock & Mueller, 2011; Raykov & Marcoulides, 2019), omega hierarchical is often preferred to other alternatives when a bi-factor model is involved (Savalei & Reise, 2019). In particular, if researchers are interested in the total score reliability, applying alternative statistics such as omega total (i.e., ω_t) could be more suitable (Revelle & Condon, 2019). In contrast with omega hierarchical, omega total represents the ratio of variance accounted for by all common factors (i.e., general plus group factors) to test variance (Revelle & Condon, 2019).

Bi-factor exploratory factor analysis and reliability

Omega hierarchical depends on the quality of the estimation of the underlying general factor model. Consequently, negatively biased general factor loadings would result in a negatively biased omega hierarchical reliability, and vice versa. As Monte Carlo studies have shown that confirmatory bi-factor modeling (i.e., BCFA) resulted in biased estimation when the model is not correctly specified (e.g., true cross-loadings are unmodeled; Morin et al., 2016), many researchers have favored the application of exploratory alternatives (i.e., BEFA). To date, most common approaches to BEFA models in the reliability literature are based on the SL transformation (Rodriguez et al. 2016a, 2016b; Zinbarg et al., 2007; Zinbarg et al., 2006). However, an SL solution is a low-rank approximation for a bi-factor structure that imposes specific linear dependencies between sets of general and group factor loadings (Waller, 2018). As the imposition of such constraints does not hold in a majority of situations, SL solutions are expected to incorrectly recover the underlying factor loadings (Abad et al., 2017; Garcia-Garzon et al.,2019; Jennrich & Bentler, 2011; Mansolf & Reise, 2016). However, the impact of the distortions introduced by approximating a bi-factor model by means of an SL solution is controversial. For example, SL has been shown to produce adequate solutions when approximating simple bi-factor models (e.g., structures not including pure markers of the general factor; Figures 15–16 in Supplementary Data; Giordano & Waller, 2019). Either way, the specific effect of applying SL in omega hierarchical estimation remains uninvestigated. To stress that pure indicators play a relevant role in bi-factor modeling, as their presence enforces that proportionality constraints are not being met for items loading in the same group factor in which the pure indicator is observed. Thus, rank-deficient algorithms could struggle to adequately recover structures presenting such deviations of the simple structure solution. Regardless of the unknown nature of their origin, pure indicators have been repeatedly observed throughout the exploratory bi-factor literature, as in the Quality of Life Dataset (Chen, West & Sousa, 2006; Abad et al., 2017; Jennrich & Bentler, 2011), the Observer Alexithymia Scale (Reise et al., 2010; Jennrich & Bentler, 2012) or the Revised NEO Personality Inventory (Chen et al., 2012; Robertson, 2019). Thus, exploring the effects of pure indicators could help to reveal the extent that their presence affects the recovery of the bi-factor model under different conditions.

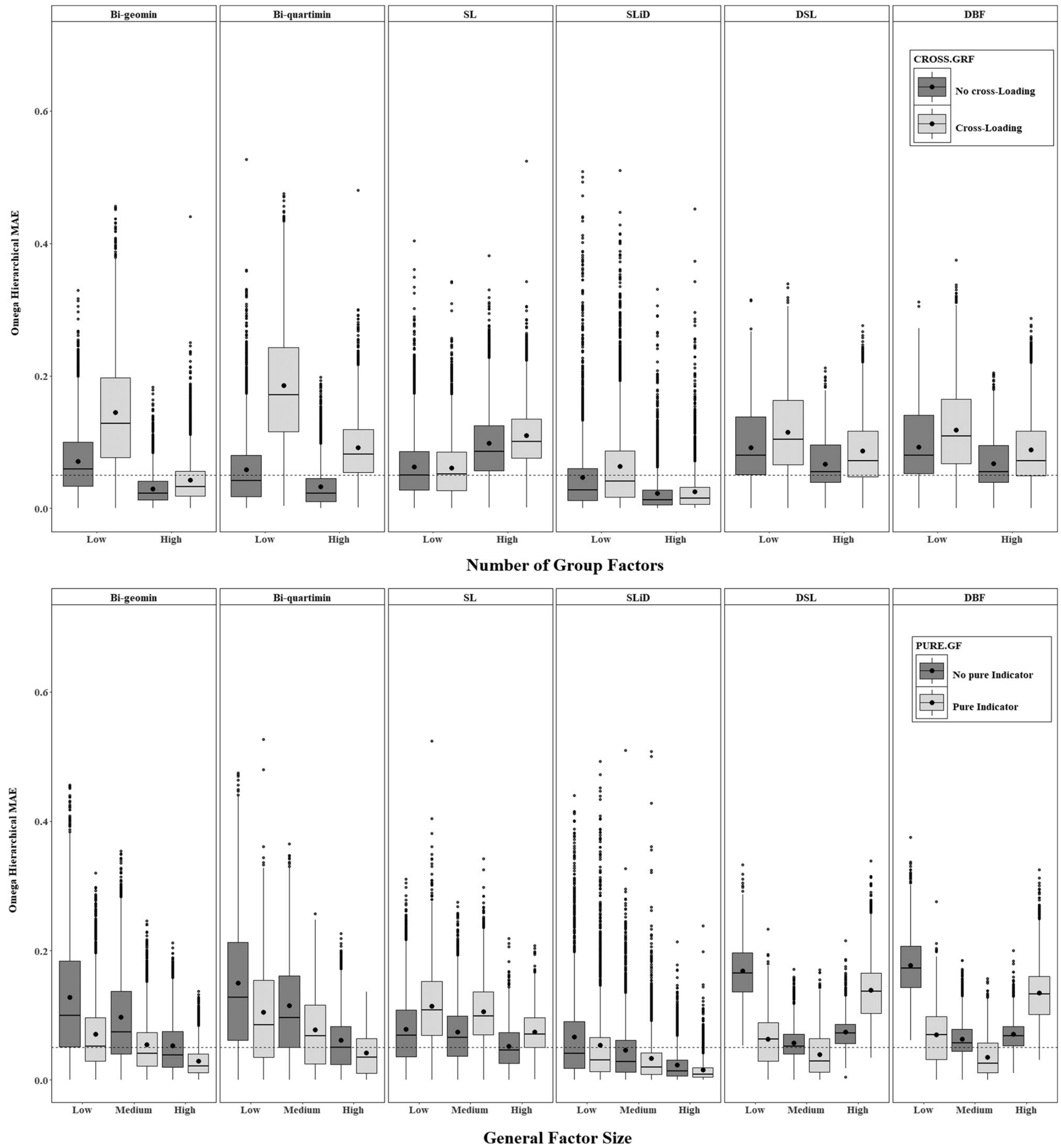

Figure 1. Omega hierarchical MAE boxplots corresponding to the two-way interactions of CROSS.GRF x NUM.GRF (upper row) and PURE.GF x SIZE.GF for bi-factor models. NUM.GRF = number of group factors; CROSS.GRF = Cross-loading presence; SIZE.GF = General factor average factor loading; MAE = Mean absolute error; SL = Schmid-Leiman orthogonalization; SLiD = Iterative Empirical Target Rotation based on an initial Schmid-Leiman solution; DSL = Direct Schmid-Leiman. DBF = Direct Bi-factor.

Alternatively, several modern algorithms for conducting BEFA have recently been proposed. These proposals may be divided into two broad families: First, two different algorithms adapting previously existent rotation criteria to recover IC (bi-quartimin) or ICB (i.e., bi-geomin) bi-factor models; and second, methods based on approximating a bi-factor exploratory model via a partially specified (i.e., empirical iterative target rotation based on Schmid-Leiman; Garcia-Garzon et al., 2019) or completely specified target rotation (i.e., the Direct Schmid-Leiman and Direct Bi-Factor algorithm; Giordano & Waller, 2019; Waller, 2018).

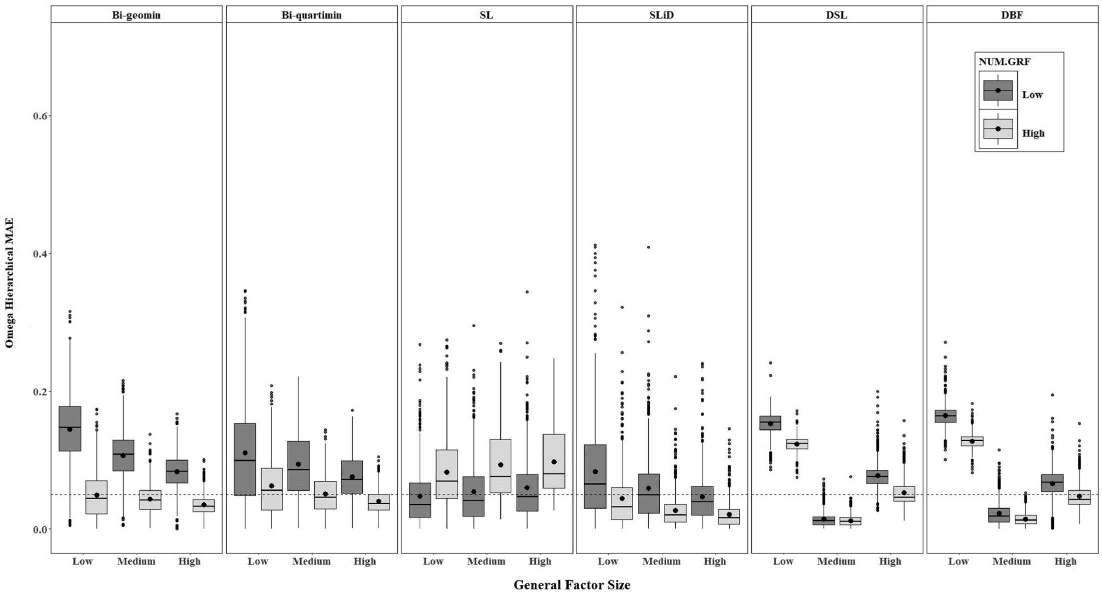

Figure 2. Omega hierarchical MAE boxplots corresponding to the two-way interaction of NUM.GRF x SIZE.GF for second-order models. NUM.GRF = number of group factors; SIZE.GF = General factor average factor loading; MAE = Mean absolute error; SL = Schmid-Leiman orthogonalization; SLiD = Iterative Empirical Target Rotation based on an initial Schmid-Leiman solution; DSL = Direct Schmid-Leiman. DBF = Direct Bi-factor.

Modern approaches to bi-factor exploratory factor analysis

Bi-geomin and bi-quartimin criteria

Jennrich and Bentler (2011, 2012) developed two bi-factor rotations widely applied in the literature. First, bi-quartimin was proposed to recover simple IC structures successfully (Jennrich & Bentler, 2011)

$$B(\Lambda) = quartimin(\Lambda_2) = \sum_{i=1}^{p}\sum_{j=2}^{k}\sum_{j'=j+1}^{k}\lambda_{ij}^2\lambda_{ij'}^2 \tag{3}$$

Later, the same authors introduced the adaptation of the geomin criterion to the bi-factor case in order to better approximate ICB structures (bi-geomin; Jennrich & Bentler, 2012)

$$B(\Lambda) = geomin(\Lambda_2) = \sum_{i=1}^{p}\prod_{j=2}^{k}(\lambda_{ij}^2 + \varepsilon)^{1/m} \tag{4}$$

where ε represents a small quantity (i.e., .01) that serves to make the function differentiable (Hattori et al., 2017).

A characteristic of both bi-factor rotation criteria is that their value is computed when rotating the group factors to a simple solution (Eqs. 3 and 4; also see Jennrich & Bentler, 2011). However, there is no guarantee that the general factor loadings will be in the manifold of acceptable solutions after this step. To ensure the appropriateness of the final solution, the complete structure (including general factor loadings) is projected into this manifold in a second (projection) step. Unfortunately, this procedure renders bi-geomin and bi-quartimin prone to shifting the variance contained in one of the group factors to the general factor and to produce factor collapse (Mansolf & Reise, 2016). Other authors have also found these methods to overestimate general factor loadings (Revelle & Wilt, 2013, p. 495). Lastly, bi-geomin is expected to be more accurate than bi-quartimin for complex structures (i.e., ICBP; Abad et al., 2017), but not for simpler structures (ICB; Figure 2; Giordano & Waller, 2019).

Bi-factor rotation via target rotation

Following a different strategy, the SLiD and the DSL-DBF algorithms are based on the flexibility of the target rotation to approximate factor solutions with a pre-defined pattern model. In both algorithms, once a target matrix following a bi-factor pattern is found, a final solution is obtained by rotating a factor loading matrix (of the same dimensions as the bi-factor model) toward the bi-factor target matrix. The main differences lie in how each method defines such a bi-factor target matrix.

The SLiD algorithm (Garcia-Garzon et al., 2019) presents three main characteristics: a) It estimates a partially specified target matrix (Browne, 2001). In this target matrix, elements expected to be negligible in the final rotated solution are given a target value of zero, so the final value of the correspondent loading is minimized during the rotation step. The remaining elements are not constrained (i.e., not considered when performing the rotation); b) Loadings to be minimized are identified by estimating an appropriate empirical cutoff point for each group factor separately, as illustrated in Appendix 1 using the Thurstone 9 mental test dataset (discussed in McDonald, 1999): First, squared, Kaiser-row normalized factor loadings for the specific factor loadings obtained from an SL solution are computed (Panel B, Table A1). Second, for each group factor, these transformed loadings are sorted by increasing value (Panel C, Table A1). Third, one-lagged loadings differences (i.e., the difference between each loading and its predecessor) are obtained (Panel D, Table A1). Lastly, the cutoff is set as the loading that falls immediately below the first difference that is greater than the average difference between all consecutive sorted loadings for that factor (illustrated in Panel C, Table A1). Thus, this cutoff aims to avoid wrongly setting to zero a target element when a relevant cross-loading is present, as occurs for the "Pedigree" item; c) SLiD, based on Moore et al. (2015), improves the obtained solution using an iterative refinement of the target matrix (Abad et al., 2017).

SLiD has been shown to recover IC, ICB, ICP, and ICBP structures presenting multiple cross-loadings and weak factors (Garcia-Garzon et al., 2019). Moreover, it outperformed SLi, a predecessor using a similar algorithm flow that applied fixed cutoff points. Notably, to the extent that SLi has been evidenced to outperform bi-geomin and bi-quartimin (Abad et al., 2017; Giordano & Waller, 2019), SLiD is also expected to provide a better reliability estimation than those methods. Lastly, as SLiD led to factor collapse in some instances, it is hypothesized that minor positive omega hierarchical bias might occur.

The DSL (i.e., Direct Schmid-Leiman) and DBF (i.e., Direct Bi-Factor) algorithms (Waller, 2018) differ from SLiD in each of the three characteristics previously examined: (a) Both DSL and DBF apply a completely specified target rotation where all elements of the target matrix are given a value toward which loadings are maximized or minimized. In this kind of target, elements associated with expected negligible factor loadings are also fixed to zero, whereas the remaining ones are set to one. Consequently, the rotation criterion depends on all elements of the target matrix; (b) to decide which loadings should be given a value of zero, all elements of an initial correlated-factor solution of a dimension less than the expected bi-factor model are compared with a single fixed cutoff point (e.g., .25). Additionally, a column vector of ones is added to the target matrix so that its dimensionality is similar to that of the final solution. The difference between DSL and DBF lies in the former rotating the original correlated factors solution plus an additional column vector of zeroes, whereas the latter rotates a correlated factors solution of the expected bi-factor dimensionality. Accordingly, DSL will result in a limited-rank solution (similarly to SL), while DBF will estimate a full-rank solution (i.e., a bi-factor model). Lastly, both algorithms compute a single iteration target rotation by means of orthogonal projection, as defined in Schönemann (1966). DSL has been shown to provide the optimal (in a least-square sense) rank-deficient approximation to a bi-factor structure if the true target given is known (Waller, 2018). Based on previous studies, DSL is expected to provide adequate results when applied to recover either type of structure (i.e., full or deficient rank solutions), and to outperform DBF under many conditions (Giordano & Waller, 2019).

Unfortunately, DSL, DBF, and SLiD have never been previously compared in the literature. It could be hypothesized that applying a completely specified target rotation would introduce error in the rotation process to such an extent that the true factor loadings differ from the given value specified in the target matrix. As loadings corresponding to freed non-zero targets are maximized to be as close as possible to one, cross-loadings or pure indicators (or other incorrect loadings targeted as one) could see their values inflated in the final solution. However, DSL has been shown to outperform SLi (an ancestor of SLiD using fixed cutoff points) under structures without pure markers of the general factor (Giordano & Waller, 2019). Accordingly, even though DSL might outperform SLiD when recovering full or limited rank structures if substantial deviations of proportionality constraints do not occur (e.g., when pure indicators of the general factor are present), further research is needed. Moreover, it is crucial to bear in mind that the translation of errors in factor loading recovery to omega hierarchical bias would depend on the number, location, and magnitude of these deviations.

In conclusion, this article has two main objectives: a) to understand the extent to which the studied algorithms can correctly recover model-based reliability; and b) to evaluate the functioning of each method when recovering bi-factor population structures with differing levels of complexity (i.e., IC, ICB, ICP, or ICBP). Additionally, we aim to investigate whether algorithms producing full rank (i.e., bi-geomin, bi-quartimin, SLiD, and DBF) or low-rank (i.e., SL and DSL) solutions perform better when recovering omega hierarchical from a matching dimensionality. Additionally, we aim to understand the extent to which each algorithm could misguide researchers into believing that a general factor is present when there is none. These results are of particular interest regarding how the selection of the BEFA algorithm could influence the debates over whether a general factor is feasible in areas such as psychopathology (Caspi & Moffitt, 2018) or personality (Revelle & Wilt, 2013). Accordingly, three different Monte Carlo simulation studies were conducted to study each type of structure (i.e., bi-factor, second-order, or structures without general factor).

Study 1: Bi-factor structures

Method

Data simulations and parameter estimations were performed in R 3.6.0 (R Core Team, 2019), with analyses of variance (ANOVA) conducted in Jamovi 1.0.0.0 (Jamovi Project, 2019).

Manipulated factors

Several variables commonly manipulated in the factor analysis literature were manipulated, namely:

- Sample size (SAMPLE): low = 150, medium = 500, high = 1000.
- Number of group factors (NUM.GRF): low = 3, high = 6.
- Number of variables per group factor (VAR.GRF): low = 4, high = 8.
- Cross-loadings on the group factors (CROSS.GRF): no or yes.
- Pure indicators of the general factor (PURE.GF): no or yes.
- General factor loading size (SIZE.GF): low = .30, medium = .45, high = .60.

The group factor loading size was not manipulated but varied across each group factor. Specifically, the number of group factors was divided into thirds, with each third presenting a different average loading size (low: $\bar{\lambda}$ = .30; medium: $\bar{\lambda}$ = .45; high: $\bar{\lambda}$ = .60). Factor loadings for the general and group factors were defined within a ± .10 range (e.g., general factor loadings for SIZE.GF = high ranged from .50 to .70). We note that manipulating the general factor size implies manipulating the magnitude of omega hierarchical.

For conditions including cross-loadings (CROSS.GRF = *yes*), an additional factor loading of .30 was added for the item with the highest group loading of each group factor in the following group factor. Item communality was held constant for these items by subtracting a small amount from its remaining factor loadings, except for the indicators of the low loading factor(s). For these items, such an amount was subtracted from the general factor. For conditions including a pure indicator (PURE.GF = *yes*), the middle item of each group factor was substituted by a factor loading of .01. Item communality was held constant by increasing the corresponding general factor loading. Thus, true omega hierarchical values would be decreased when introducing cross-loadings and increased when introducing pure indicators. As conditions were fully crossed, a total of $2 \times 2 \times 2 \times 2 \times 2 \times 3 = 144$ conditions were studied. Additionally, the combination of cross-loadings (CROSS.GRF) and pure indicators (PURE.GF) led to the structure typology (STRUCTURE) previously mentioned: IC, ICB, ICP, and ICBP. This taxonomy has been shown to be useful when studying differences across methods (Abad et al., 2017).

Data simulation

A total of 100 sample correlation matrices were simulated for each condition. For each replication, general factor loadings were first randomly sorted to avoid proportionality constraint effects. The population correlation matrices were specified from these population factor-loading matrices by inserting unities in the diagonal of the reproduced population correlation matrix—afterwards, sample scores matrices of SAMPLE × NUM.GRF × VAR.GRF dimensions were obtained from random standard normal deviates simulated with the package *mvtnorm* (Genz et al., 2017). For all solutions, group factors were aligned with their corresponding population factor structure following the least-squares criterion using the *faAlign* function (Waller, 2019).

Statistical analyses

Six BEFA algorithms were scrutinized: Bi-quartimin, bi-geomin, SL, SLiD, DSL, and DBF. The SL and SLiD

Table 1. Marginal omega hierarchical mean absolute error (MAE) and mean bias error (MBE) for each method for bi-factor structures.

Variable / Level	Bi-geomin	Bi-quartimin	SL	SLiD	DSL	DBF
SAMPLE						
150	.088	.109	.118	**.061**	.094	.096
	(.085)	(.104)	(−.112)	(−.007)	(−.003)	(.006)
500	.067	.089	.073	**.033**	.089	.090
	(.067)	(.085)	(−.069)	(.007)	(.009)	(.015)
1000	.061	.079	.058	**.025**	.087	.089
	(.061)	(.077)	(−.054)	(.008)	(.010)	(.015)
NUM.GRF						
3	.108	.122	.062	**.055**	.103	.105
	(.107)	(.117)	(−.052)	(.011)	(.004)	(.013)
6	.036	.062	.104	**.024**	.077	.078
	(.035)	(.061)	(−.104)	(−.004)	(.007)	(.011)
VAR.GRF						
4	.077	.096	.085	**.048**	.095	.097
	(.075)	(.092)	(−.076)	(.008)	(.007)	(.002)
8	.067	.088	.083	**.031**	.085	.087
	(.067)	(.089)	(−.081)	(−.002)	(.018)	(.022)
CROSS.GRF						
No	.050	.045	.081	**.035**	.079	.080
	(.048)	(.039)	(−.080)	(−.006)	(−.005)	(.000)
Yes	.094	.138	.085	**.044**	.101	.103
	(.093)	(.138)	(−.077)	(.012)	(.016)	(.024)
PURE.GF						
No	.093	.109	.068	**.045**	.100	.104
	(.091)	(.107)	(−.059)	(.013)	(.050)	(.056)
Yes	.052	.075	.098	**.034**	.080	.080
	(.050)	(.070)	(−.097)	(−.007)	(−.039)	(−.033)
SIZE.GF						
Low	.100	.127	.095	**.060**	.116	.123
	(.097)	(.121)	(−.086)	(−.007)	(.111)	(.120)
Medium	.076	.097	.090	**.040**	.048	.049
	(.075)	(.097)	(−.086)	(.008)	(.012)	(.018)
High	.041	.052	.063	**.019**	.106	.102
	(.041)	(.051)	(−.063)	(.008)	(−.106)	(−.102)
STRUCTURE						
IC	.064	.059	.068	**.039**	.084	.088
	(.063)	(.055)	(−.066)	(.000)	(.039)	(.043)
ICB	.121	.159	.069	**.051**	.116	.120
	(.120)	(.159)	(−.052)	(.025)	(.062)	(.070)
ICP	.036	.032	.093	**.030**	.074	.072
	(.034)	(.034)	(−.093)	(−.013)	(−.079)	(−.043)
ICBP	.067	.118	.103	**.038**	.086	.087
	(.066)	(.117)	(−.103)	(−.000)	(−.030)	(−.022)
AVERAGE	.072	.092	.083	**.040**	.090	.092
	(.071)	(.089)	(−.078)	(.003)	(.005)	(.012)

Note. SAMPLE = sample size; NUM.GRF = number of group factors; VAR.GRF = number of indicators per group factor; CROSS.GRF = Cross-loading presence; PURE.GF = Pure indicator presence; SIZE.GF = General factor average factor loading; IC = Independent Cluster; ICB = Independent Cluster Basis; ICP = Independent Cluster with Pure Indicators; ICBP = Independent Cluster Basis with Pure Indicators; SL = Schmid-Leiman orthogonalization; SLiD = Iterative Empirical Target Rotation based on an initial Schmid-Leiman solution; DSL = Direct Schmid-Leiman. DBF = Direct Bi-factor. MBE values appear under brackets. Conditions with MAE ≥ .05 appear shadowed in gray. Lowest MAE for each condition appears bolded.

Table 2. Univariate Analysis of Variance (ANOVA) effect sizes for the mean absolute error (MAE) across methods for bi-factor structures.

Effect Type / Variables	Bi-geomin	Bi-quartimin	SL	SLiD	DSL	DBF
Study 1: Bi-factor models						
SAMPLE	.101	.105	.350	.127	.012	.014
NUM.GRF	.518	.394	.270	.130	.202	.213
VAR.GRF	.019	.013	.000	.042	.031	.034
CROSS.GRF	.282	.612	.005	.014	.148	.161
PURE.GF	.261	.177	.161	.017	.148	.168
SIZE.GF	.323	.412	.151	.148	.569	.579
NUM.GRF * CROSS.GRF	.159	.173	.010	.007	.001	.002
SIZE.GF* PURE.GF	.037	.020	.007	.001	.640	.633
Study 2. Second-order models						
SAMPLE	.053	.271	.530	.101	.134	.123
NUM.GRF	.525	.234	.241	.133	.344	.335
VAR.GRF	.003	.005	.001	.034	.024	.005
SIZE.SOF	.184	.089	.028	.084	.638	.628
NUM.GRF * SIZE. SOF	.083	.005	.003	.009	.179	.142
Study 3: No general factor						
SAMPLE	.001	.028	.007	.056	.002	.075
NUM.GRF	.583	.161	.215	.118	.590	.525
VAR.GRF	.117	.002	.000	.068	.753	.700
CROSS.GRF	.046	.365	.142	.059	.073	.051

Note. SAMPLE = sample size; NUM.GRF = number of group factors; VAR.GRF = number of indicators per group factor; CROSS.GRF = Cross-loading presence; PURE.GF = Pure indicator presence; SIZE.GF = General factor average factor loading; SIZE.SOF = Second-order general factor average factor loading; SL = Schmid-Leiman orthogonalization; SLiD = Iterative Empirical Target Rotation based on an initial Schmid-Leiman solution; DSL = Direct Schmid-Leiman. DBF = Direct Bi-factor. The dependent variable in the ANOVAs was omega hierarchical MAE. All main effects and interactions presenting a large effect (η_p^2>.14) are shown, with large (η_p^2 >.14) effects in gray shadow.

algorithms were computed using the code in Garcia-Garzon et al. (2019). The initial correlated factor solution for SL was obtained using oblimin rotation ($\gamma = 0$). DSL and DBF algorithms were computed using the *fungible* package (Waller, 2019), with the default cutoff point of .25 for the target matrix definition. The correlated factor solution was obtained using the same rotation as in SL. Bi-quartimin was computed using the routine defined within the *GPArotation* package (Bernaards & Jennrich, 2005). An orthogonal bi-geomin criterion was defined (code available on request) with $\varepsilon = .01$ (following recommendations in Hattori et al., 2017) and applied using the GForth routine within the *GPArotation* package (Bernaards & Jennrich, 2005). All algorithms based on *GPArotation* (SLiD, bi-geomin, and bi-quartimin) were fitted using 100 different orthogonal random starts and 5000 maximum iterations. In all cases, factor estimation was conducted using the unweighted least squares estimator (ULS).

Lastly, ANOVAs were conducted to determine which conditions most affected the recovery of omega hierarchical. From these analyses, the partial eta squared (η_p^2) effect sizes were reported following Cohen's (1988) guidelines such that effect sizes of $\eta_p^2 > .01$, $\eta_p^2 > .06$, and $\eta_p^2 > .14$ were considered small, medium, and large effects, respectively.

Dependent variables

Omega hierarchical was computed for all population structures and solutions estimated. Two different main measures of accuracy were explored, the mean absolute error (MAE) and the mean bias error (MBE). Both represent the difference between the estimated and the population omega hierarchical values averaged across replicates within the same condition. The former considers differences in absolute value to avoid error suppression if differences of opposite signs within the same condition are averaged. Therefore, it is insensitive to the direction of the differences. The latter (i.e., MBE) represents the opposite situation, with differences computed considering the sign of the omega hierarchical values. Consequently, MAE was preferred to understand factors behind estimation errors, and MBE to explore potential patterns of omega hierarchical due to under- and overestimation for each method. As a heuristic, MAE and MBE rates over .05 were considered substantial. As recommended by a reviewer, the root mean square error (RMSE) of estimation was also explored. The RMSE considers the square root of mean differences of squared errors between population and estimated solutions per condition. As MAE and RMSE results were in the same direction, marginal RMSE rates are included in Appendix 2.

Results

Marginal MAE and MBE values are reported in Table 1. Invalid solutions (i.e., non-convergent or Heywood cases) rates ranged from .04% (SL and SLiD) to .76% (bi-geomin) of the cases. DSL and DBF provided a valid solution in all cases. Overall, the results suggest that: a) SLiD was the most accurate method under most conditions; b) Pure indicators had a positive effect on omega hierarchical recovery for all methods but SL; c) Increased general factor size was beneficial for all methods but DSL and DBF; d) No algorithm provided an adequate recovery (MAE $<$ .05) under conditions of small samples, low numbers of factors, low general factor sizes, or ICB structures.

In general, SLiD was the most accurate algorithm (MAE [SLID] = .040). SLiD showed its best performance for structures presenting a strong general factor or including pure indicators (i.e., MAE $<$ .05). Even when SLiD provided a subpar performance under certain conditions (e.g., MAE [SAMPLE = 150] = .061), it continued outperforming all alternative methods. For SLiD, MBE rates were substantially lower than MAE, positive, and close to zero, suggesting that this algorithm provided an unbiased estimation under most conditions (MBE [SLiD] = .003).

On the contrary, SL did not produce an adequate performance under any condition (i.e., MAE [SL] $<$ .05). As expected, pure indicators severely affected SL. Notably, SL was the only method to present a systematic negative bias in general (MBE [SL] = −.078) and across all conditions. DSL and DBF presented a similar functioning (with largest MAE difference between DSL-DBF: [SIZE.GF = Low] = .007]), where the latter almost never outperformed the former. Remarkably, both algorithms recovered structures with medium-sized general factors approximately as well as SLiD but failed to provide an adequate omega estimation otherwise. MBE results highlighted that under alternative SIZE.GF conditions, both methods substantially overestimated/underestimated omega hierarchical when a low/high general factor was present, respectively.

Lastly, the results support previous hypotheses regarding the functioning of bi-quartimin and bi-geomin. First, bi-quartimin presented a worse overall performance than bi-geomin (MAE ([bi-quartimin] = .092; MAE [bi-geomin] = .072). Second, both methods' performance was substantially hampered by the presence of cross-loadings, particularly if no pure indicators were present (i.e., ICB structures). Lastly, MBE results suggest that both methods routinely overestimated omega hierarchical values across all conditions (MBE ([bi-quartimin] = .071; MBE [bi-geomin] = .089).

As reflected in Table 2, SLiD was shown to be the most robust method (i.e., its performance was the least affected by the manipulated variables). Nevertheless, the general factor size was a relevant condition for all methods (η^2_p [SIZE.GF] = .148 $\leq \eta^2_p \leq$ .579), where increasing general factor size was beneficial for all methods but DSL and DBF (Table 2). Additionally, all methods presented a medium to strong effect of the number of factors (η^2_p [NUM.GRF] =.130 $\leq \eta^2_p \leq$ .518), where increasing the number of factors boosted recovery for all methods but SL (as shown in Table 2). The presence of pure indicators strongly affected all methods (η^2_p [PURE.GF] = .148 $\leq \eta^2_p \leq$ .261) except SLiD (η^2_p [PURE.GF] = .017), such that their presence improved omega hierarchical recovery for all methods but SL (Table 2).

Two relevant interactions were noted. In the case of bi-geomin and bi-quartimin, the estimation of structures with three factors exacerbated the negative effect of cross-loadings (bi-quartimin: η^2_p [NUM.GRF * CROSS. SF] = .173 bi-geomin: η^2_p [NUM.SF * CROSS. SF] = .159; Top row; Figure 1; Table 2). For DSL and DBF, a strong interaction between general factor size and the presence of pure indicators was observed (DSL: η^2_p [PURE.GF * SIZE.GF] = .640; DBF: η^2_p [PURE.GF * SIZE.GF] = .633; Bottom row;

Table 3. Marginal omega hierarchical mean absolute error (MAE) and mean bias error (MBE) across each method for second-order models.

Variable / Level	Bi-geomin	Bi-quartimin	SL	SLiD	DSL	DBF
SAMPLE						
150	.088	.105	.122	**.066**	.079	.082
	(.084)	(.099)	(−.116)	(−.020)	(.033)	(.039)
500	.073	.065	.061	**.041**	.070	.071
	(.072)	(.061)	(−.058)	(.011)	(.030)	(.035)
1000	.070	.049	.038	**.034**	.068	.069
	(.070)	(.046)	(−.035)	(.007)	(.021)	(.032)
NUM.GRF						
3	.112	.093	**.054**	.063	.082	.085
	(.110)	(.088)	(−.046)	(.007)	(.029)	(.040)
6	.043	.051	.091	**.031**	.063	.064
	(.040)	(.049)	(−.091)	(−.008)	(.027)	(.031)
VAR.GRF						
4	.079	.070	.075	**.055**	.070	.073
	(.078)	(.099)	(−.046)	(.007)	(.029)	(.040)
8	.075	.075	.072	**.039**	.074	.073
	(.073)	(.071)	(−.070)	(−.003)	(.028)	(.032)
SIZE.SOF						
Low	.097	.087	.066	**.064**	.138	.146
	(.092)	(.076)	(−.056)	(−.014)	(.138)	(.146)
Medium	.075	.073	.074	.043	**.013**	.019
	(.074)	(.072)	(−.072)	(.006)	(.009)	(.017)
High	.059	.058	.079	**.034**	.065	.057
	(.059)	(.058)	(−.079)	(.006)	(−.065)	(−.057)
AVERAGE	.077	.072	.073	**.047**	.072	.074
	(.075)	(.068)	(−.069)	(.001)	. (028)	(.035)

Note. SAMPLE = sample size; VAR.GRF = number of indicators per group factor; NUM.GRF = number of group factors; SIZE.SOF = Second-order general factor average factor loading; SL = Schmid-Leiman orthogonalization; SLiD = Iterative Empirical Target Rotation based on an initial Schmid-Leiman solution; DSL = Direct Schmid-Leiman. DBF = Direct Bi-factor. MBE values appear under brackets. Conditions with MAE ≥ .05 appear shadowed in gray. Lowest MAE for each condition appears bolded.

Figure 1). Specifically, while the presence of pure indicators improved omega recovery if the general factor was low, it hampered omega estimation when the general factor size was high. Lastly, as reflected in Figure 1, it was remarkable that SLiD and DBF/DSL presented the largest and lowest variability in their estimation. This noticeable number of undesirable solutions in the former algorithm is further explored in the Discussion section.

Study 2: Second-order structures

Method

Manipulated factors. A total of four relevant variables were manipulated in Study 2:

- Sample size (SAMPLE): low = 150, medium = 500, high = 1000.
- Number of group factors (NUM.GRF): low = 3, high = 6.
- Number of variables per group factor (VAR.GRF): low = 4, high = 8.
- Second-order factor loading size (SIZE.SOF): low = .55, medium = .70, high = .80.

As can be seen, the sample size, number of group factors, and number of variables per group factor conditions were the same as those of Study 1. However, neither cross-loadings nor pure indicators were investigated. Second-order structures were simulated to ensure that the SIZE.SOF conditions were analogous to the SIZE.GF conditions of Study 1. Also, as in Study 1, each third of the group (first-order) factors presented a different average loading size (low: $\bar{\lambda}$ = .40; medium: $\bar{\lambda}$ = .55; high: $\bar{\lambda}$ = .70). Factor loadings on the group factors were defined within a ± .10 range, while loadings on the second-order factor were all of equal magnitude within conditions.

Once the second-order structures were generated, the function sim.hierarchical from the psych package (Revelle, 2018) was used to generate population correlation matrices, with population omega hierarchical values obtained from a Schmid-Leiman transformation. Such SL solutions are ensured to reproduce population correlation matrices with MAE = 0. The data simulation, statistical analysis, and dependent variable specifications were those of Study 1. RMSE results are again presented in Appendix 2.

Results

Marginal MAE values are reported in Table 3. Invalid solutions (i.e., non-convergent or Heywood cases) rates ranged from .01% (bi-geomin and bi-quartimin) to 2.03% (SLiD) of the cases. DSL and DBF provided a valid solution in all cases. Surprisingly, SLiD recovered omega hierarchical better than any alternative method (MAE [SLiD] = .047), with alternative methods providing a similar overall performance (e.g., .072 ≤ MAE ≤ .077 for the remaining methods). Overall, all the manipulated variables produced similar effects to those observed for bi-factor solutions, with notable exceptions commented on below.

Even though SL outperformed alternative algorithms when the number of factors was low (MAE [NUM.GRF = 3] = .054), it systematically resulted in negative bias for all conditions (MBE [SL] = −.069). For second-order structures, bi-quartimin slightly outperformed bi-geomin (MAE [Bi-quartimin] = .072; MAE [Bi-geomin] = .077). However, such an effect could be associated with the lack of cross-loadings or pure indicators in the simulated second-order models. As observed in Study 1, both methods: a) Failed to show an adequate recovery of omega hierarchical in a

Table 4. Marginal omega hierarchical mean absolute error (MAE) across each method for models without a general factor.

Variable / Level	Bi-geomin	Bi-quartimin	SL	SLiD	DSL	DBF
SAMPLE						
150	.257	.214	.149	**.092**	.499	.528
500	.255	.172	.136	**.054**	.497	.513
1000	.250	.162	.131	**.047**	.496	.509
NUM.GRF						
3	.348	.241	.093	**.094**	.462	.486
6	.160	.124	.184	**.035**	.533	.547
VAR.GRF						
4	.225	.188	.137	**.086**	.446	.472
8	.283	.177	.140	**.043**	.549	.560
CROSS.GRF						
No	.236	.082	.104	**.045**	.486	.510
Yes	.271	.284	.174	**.085**	.506	.523
AVERAGE	.254	.183	.138	**.064**	.497	.516

Note. SAMPLE = sample size; NUM.GRF = number of group factors; VAR.GRF = number of indicators per group factor; CROSS.GRF = Cross-loading presence; SL = Schmid-Leiman orthogonalization; SLiD = Iterative Empirical Target Rotation based on an initial Schmid-Leiman solution; DSL = Direct Schmid-Leiman. DBF = Direct Bi-factor. MAE and MBE values are equal given that true omega hierarchical is zero in all conditions. -. Conditions with MAE ≥ .05 appear shadowed in gray. Lowest MAE for each condition appears bolded.

majority of occasions; and b) resulted in positively biased estimation of omega hierarchical under all conditions (MBE [Bi-quartimin] = .075; MBE [Bi-geomin] = .078). Lastly, DSL and DBF presented similar functioning in Studies 1 and 2. Depending on whether the general factor was low/high in magnitude, DSL overestimated/underestimated omega hierarchical, respectively. As before, DBF performance was similar to or worse than that of DSL for almost all conditions. Nevertheless, DSL showed the best recovery of omega hierarchical if the general factor was of medium size (MAE [SIZE.GF = Medium] = .013).

As in Study 1, SLiD was the most robust algorithm (see Table 2). The number of factors was decisive for all methods (η_p^2 [NUM.GRF] =.133 $\leq \eta_p^2 \leq$ .525). As with bi-factor structures, increasing the number of factors reduced MAE for all methods but SL. Contrary to Study 1, SIZE.SOF only presented a strong effect for bi-geomin, DSL, and DBF (η_p^2 [SIZE.GF] = .184 $\leq \eta_p^2 \leq$ .638). Lastly, the sample size was mostly relevant for bi-quartimin and SL (η_p^2 [SAMPLE] = .271 $\leq \eta_p^2 \leq$ .530). Additionally, a two-way interaction between the number of group factors and second-order factor size was observed for DSL and DBF (DSL: η_p^2 [NUM.GRF * SIZE.SOF] = .179; DBF: η_p^2 [NUM.GRF * SIZE.SOF] = .142; Figure 2), where structures with three factors were more incorrectly estimated when the general factor was either low or high. As also reflected in Figure 2, SLiD presented several outliers corresponding to inadequate solutions. Again, DSL and DBF resulted in the most consistent estimation across all methods.

Study 3: Structures without a general factor

Method

Manipulated factors

In this Monte Carlo study, a new simulation with a no-general-factor condition was studied. That is, all general factor loadings were zero, reproducing a condition where no general factor exists. A total of four relevant variables were manipulated in the same manner as Study 1:

- Sample size (SAMPLE): low = 150, medium = 500, high = 1000.
- Number of group factors (NUM.GRF): low = 3, high = 6.
- Number of variables per group factor (VAR.GRF): low = 4, high = 8.
- Cross-loadings on the group factors (CROSS.GRF): no or yes.

Pure indicators of the general factor were not introduced in this study, as no variables loaded on the general factor. The data simulation, statistical analyses, and dependent variable specifications were those of Study 1. RMSE results are again presented in Appendix 2. As the true omega hierarchical was zero in all cases and the numerator in Eq. 2 was always positive, MAE and MBE were equivalent in this study.

Results

MAE results are presented in Table 4 (with MAE and MBE values equal, given that true omega hierarchical is zero under all conditions). Low invalid solutions rates (i.e., non-convergent or Heywood cases) were observed for all methods but DSL and DBF: .01% for bi-geomin and bi-quartimin; .02% for SL; .05% for SLiD). Overall, no algorithm provided an overall accurate recovery of omega hierarchical recovery when no general factor was present. Even though SLiD presented adequate performance under certain conditions (e.g., large sample size, many group factors), its performance was on the whole unsatisfactory (MAE[SLiD] = .064). Nevertheless, the performance of SLiD was still considerably better than the alternatives, which produced large (MAE[SL] = .138) to extremely large (MAE[DBF] = .516) levels of overall error. SL overestimated omega hierarchical under all conditions (best performance: MAE [NUM.GRF = Low] = .093). Of the remaining algorithms, only bi-quartimin under the no cross-loading condition resulted in MAE rates under .10 (MAE

Table 5. Estimated Omega hierarchical for all studied algorithms for seven classic bi-factor examples and a second-order structure.

	Structure	Bi-geomin	Bi-quartimin	SL	SLiD	DSL	DBF
Thurstone (3)	ICB	.83	.82	.74	.77	.66	.66
Thurstone/Bechtoldt (6)	ICB	.80	.81	.72	.78	.80	.80
Holzinger & Swineford (4)	ICBP	.81	.83	.64	.79	.68	.69
Brigham/Thurstone (3)	ICBP	.90	.89	.85	.80	.58	.59
Harman/Holzinger (5)	ICBP	.87	.88	.66	.78	.63	.64
Reise, Morizot & Hays (5)	ICB	.87	.88	.82	.84	.77	.77
Chen, West & Souza (4)	ICBP	.85	.86	.78	.84	.76	.76
Jensen & Weng (3) *	IC	.73	.72	.69	.72	.60	.60
AVERAGE		.83	.84	.74	.79	.68	.68

Note. * Second-order structure. SL = Schmid-Leiman orthogonalization; SLiD = Iterative Empirical Target Rotation based on an initial Schmid-Leiman solution; DSL = Direct Schmid-Leiman. DBF = Direct Bi-factor. DSL and DBF algorithms computed with default cutoff point (.25).

[CROSS.GRF = No] = .082). Bi-geomin, DSL, and DBF resulted in a severe overestimation of the non-present general factor under all marginal conditions studied.

The number of factors was revealed to be a highly influential factor when recovering omega hierarchical in no-general factor structures for all methods studied (η^2_p [NUM.GRF] = .161 ≤ η^2_p ≤ .590) except SLiD (η^2_p [SIZE.GF] = .118), for which it produced a medium-size effect (see Table 2). For bi-geomin, bi-quartimin, SL, and SLiD, increasing the number of factors reduced omega overestimation, while the opposite was true for DSL and DBF. Moreover, the presence of cross-loadings was particularly detrimental for bi-quartimin (η^2_p [CROSS.GRF = yes] = .365) and SL (η^2_p [CROSS.GRF = yes] = .142). Lastly, increasing the number of variables negatively impacted the estimation of omega hierarchical for DSL (η^2_p [VAR.GRF] = .753) and DBF (η^2_p [VAR.GRF] = .700).

An analysis of eight examples

The performance of each algorithm was further compared over a set of eight datasets traditionally considered to follow a bi-factor (first seven datasets) or second-order structure (last example): Thurstone's nine mental tests (discussed in McDonald, 1999); Thurstone and Bechtold's 17 mental tests (Bechtoldt, 1961); Holzinger and Swineford's 14 tests (Holzinger & Swineford, 1937); Brigham's nine tests (Thurstone, 1933); Harman's 24 mental tests (Harman, 1967); Reise, Morizot, and Hays' Consumer Assessment of Health Care Providers and Systems dataset (Reise et al., 2007); Chen, West, and Souza's Quality of Life dataset (Chen et al., 2006); and the Jensen and Weng (1994) dataset. The first five datasets represent cognitive tests for which omega hierarchical has been previously explored by means of the Schmid-Leiman transformation (Revelle & Wilt, 2013). The next two datasets were selected to characterize bi-factor assessments outside of cognitive testing. Additionally, the recovery of the Quality of Life Dataset for bi-geomin, bi-quartimin, SL procedure, and SLi-based methods has been previously discussed in the literature (Abad et al., 2017). Lastly, the Jensen and Weng (1994) set is a well-known artificial dataset (presenting a true ω_H = .69) constructed to illustrate a second-order IC model similar to those simulated in Study 2. All datasets are available in the *psych* package (Revelle, 2018). Code to reproduce these analyses is available in the Supplementary Data.

Table 5 presents the omega hierarchical estimation for each studied method. In summary, when compared with either SL or SLiD and for all cases under study, DSL and DBF tended to produce similar or lower omega hierarchical estimates. Conversely, bi-quartimin and bi-geomin consistently provided omega hierarchical estimates that were higher than any other method. Therefore, results from the analysis of these seven empirical datasets appear to mirror those obtained from Study 1. As a side note, and when compared with SL, SLiD tended to present similar or larger estimations of omega hierarchical, mainly when SL resulted in low-reliability estimates (e.g., the Harman/Holzinger dataset: SL = .66; SLiD = .79). The effect of SL presenting lower general factor loadings has also been exemplified elsewhere for the Quality of Life dataset (Abad et al., 2017), as well as for alternative ICBP and ICB constructed examples (Table 2; Mansolf & Reise, 2016). When considering the only second-order structure studied (Jensen & Weng, 1994), SL correctly recovered the true ω_H (.69), with SLiD moderately overestimating (.72) and DSL-DBF considerably underestimating (.60) the true omega hierarchical value, respectively. This dataset illustrates that under certain conditions (low number of factors, simple second-order model), SL can yield a correct solution.

Discussion

General factor modeling constitutes today a growing area of research. As bi-factor modeling has become standard practice in psychological assessment, model-based reliability estimators have also grown in popularity, with omega hierarchical playing a primary role in scale validation (Viladrich et al., 2017; Zinbarg & Alden, 2015). While omega hierarchical is commonly approached by either confirmatory modeling or the application of the SL transformation (Revelle &

Zinbarg, 2009; Zinbarg & Alden, 2015; Zinbarg et al., 2007, 2006), these strategies present substantial limitations when pure indicators (and other deviations of a simple structure) are present (Abad et al., 2017; Garcia-Garzon et al., 2019).

Accordingly, appropriate methods for directly estimating exploratory bi-factor models (BEFA) have recently been proposed in the literature: bi-geomin and bi-quartimin (Jennrich & Bentler, 2011, 2012), SLiD (Abad et al., 2017), and DSL and DBF (Waller, 2018). However, their usefulness for estimating omega hierarchical had heretofore remained unexplored. This study is the first to present evidence regarding how modern BEFA methods approximate omega hierarchical through three simulation studies and the re-analysis of eight empirical datasets.

Main remarks

First and foremost, SLiD showed the most accurate estimation of omega hierarchical across most studied conditions and structures types (either bi-factor or second-order structures). As expected, SL resulted in unsatisfactory omega hierarchical estimation for all bi-factor models studied. Moreover, SL only recovered omega hierarchical accurately for second-order models and under certain conditions (i.e., simple structures with a low number of factors, as also exemplified in the Jensen and Weng (1994) dataset). More notably, SL resulted in negative bias estimation under almost all conditions studied, which will be explained in detail below. DSL and DBF presented an adequate performance under one set of conditions: When the average general factor loadings matched those of the group factors (i.e., medium general factor size). Lastly, bi-geomin and bi-quartimin resulted in positively biased estimation of omega hierarchical on all occasions, with the latter presenting a stronger bias than the former in bi-factor models, and the reverse is true for second-order models.

These results also evidence that omega hierarchical recovery is strongly tied to the type of structure simulated, the number of factors, and the general factor size. Overall, reliability estimation benefited from larger sample sizes, a higher number of factors (but not for SL), and more items per factor. Unfortunately, no method provided a reliable recovery of omega under low sample sizes, a low number of factors, or small general factor size. In this sense, it is more challenging to accurately recover lower omega hierarchical values than higher ones. Nonetheless, the problems in recovering weak factors in bi-factor modeling have been previously shown in the literature (Garcia-Garzon et al., 2019).

Moreover, the results of Study 3 highlight the fact that the mindless application of BEFA in a case where a general factor does not exist could result in obtaining bi-factor structures with spurious, non-reliable general factors. To clarify that as ω_H is bounded below by zero, ω_H must always be positively biased under this condition. Nevertheless, even though that effect was especially pronounced for DSL and DBF algorithms, no alternative method systematically performed well. Only SLiD (and then under limited circumstances such as high sample size or a high number of factors) provided a satisfactory recovery of omega hierarchical if no general factor existed. In this sense, researchers should consider following Morin et al.'s (2016) guidelines for conducting bi-factor modeling in the context of bi-factor exploratory structural equation modeling: a) Comparing the plausibility of confirmatory and exploratory first-order correlated-factor solutions; and b) comparing this first-order solution against bi-factor and second-order solutions following confirmatory and exploratory approaches. As suggested by the authors, "this second step should be only conducted when substantive theory and the results of the first step suggest that this second source of construct-relevant multidimensionality might be present in the instrument" (Morin et al., 2016, p. 135).

Lastly, general factor size, sample size, or the number of group factors had a similar impact on the different BEFA methods under bi-factor or second-order structures. These results suggest that BEFA algorithms resulting in full-rank (i.e., SLiD) or deficient-rank solutions (i.e., DSL) could provide approximate omega hierarchical accurately under certain conditions (specific for each algorithm), regardless of the true underlying structure (as also evidenced by Giordano & Waller, 2019).

On BEFA methods for estimating general factor reliability

The SLiD algorithm

From the existent alternatives for performing direct BEFA estimation, the SLiD method was shown to be the most robust algorithm, consistently outperforming every other procedure. This study suggests that SLiD is exceptionally accurate when recovering omega hierarchical in either bi-factor or second-order structures, especially those including pure indicators of the general factor, regardless of the presence of cross-loadings. Also, SLiD was the only algorithm to provide correct (i.e., near-zero) omega hierarchical estimates under certain conditions when no general factor was present. Lastly, it should be noted that on a number of occasions, SLiD led to solutions showing a distorted factor pattern (reflected in outliers presents in Figure 1 and 2 for SLiD). Particularly, certain items (often those loading onto the weaker group factor) showed an erroneous factor pattern combining near-zero and inflated loadings along with several cross-loadings in the other group factors. Given previous warnings regarding potential factor collapse issues for SLiD (Garcia-Garzon et al., 2019; Robertson, 2019) and SLi (Giordano & Waller, 2019), researchers applying SLiD must conduct a careful examination of the estimated solutions in order to detect and discard possible distorted solutions.

Schmid-Leiman

As expected, Schmid-Leiman did not provide an adequate estimation of omega hierarchical for bi-factor models. Nevertheless, it is also true that SL presented the second-best performance after SLiD in multiple occasions, particularly when sample sizes were large or when the number of factors was low. The SL performance for second-order structures was more surprising, as SL continued to present a systematic negative bias for these structures. A detailed examination of the estimated factor loading structures revealed that SL omega hierarchical recovery was closely tied to the estimation of the first-order correlation matrix. When a general factor was present (even if it was of low strength), SL underestimated the first-order factor correlation matrix and its corresponding second-order factor loadings vector. As this underestimation was systematic, the numerator in the equation for omega (i.e., the squared sum of general factor loadings) was diminished and omega hierarchical thus underestimated. Only when a general factor was not present (and the true first-order factor correlation matrix was an identity matrix) did the opposite occur and SL overestimated the true omega hierarchical. Thus, researchers must be aware that small deviations in this first-order correlation matrix can severely impact the quality of the general factor estimation for SL. In this sense, even though oblimin was chosen because it was reported to satisfactorily recover the inter-factor correlation matrix in previous studies (Schmitt & Sass, 2011), the effects of applying different rotation criteria in SL are not always straightforward (Mansolf & Reise, 2016). Future research should address the role and effect of different first-order rotation criteria in SL estimation.

The DSL and DBF algorithms. DSL and DBF presented a subpar performance recovering omega hierarchical except when the average general factor size was of a magnitude similar to the average group factor loadings (i.e., $\overline{\lambda} = .45$). Such behavior and the similar results observed for DSL and DBF can be directly attributed to the use of completely specified target rotation. Regardless of the underlying model to be estimated, this type of target will aim to maximize a few targeted factor loadings (with are given target values of ones), with the remaining elements set as elements to be minimized (and given targets of zeroes). In this type of target, when factor loadings differ in their true value (e.g., imagine a simple target aiming to maximize two loadings of .60 and .30), and as communality must remain constant for the transformation matrix to be a proper rotation matrix, targeted values will tend to a common high value. In the case where the actual general factor loadings are more substantial than their correspondent group factor loading(s) (as in SIZE.GF = High condition), using a completely specified target rotation would always result in an increase of the latter in the expense of the former, or vice versa.

Nevertheless, it is crucial to bear in mind that the consequences of the aforementioned errors could be vastly different depending on which aspects of the factor solution are explored. For example, based on Giordano and Waller's (2019) results, DSL and DBF are expected to provide an accurate estimation of factor solutions and to function as well as SLiD for simple (i.e., without pure indicators) bi-factor structures. However, due to the systematic nature of the deviations introduced by the complete target rotation (as all targeted values have values of either one or zero), statistics based on ratios of loadings and other factor solution estimates (i.e., such as omega hierarchical) would be more severely impacted. This tradeoff could be highly relevant, as different researchers focus on different aspects of their analyses, and this issue should be explored deeply in the future.

Bi-quartimin and bi-geomin. The results confirmed two main hypotheses regarding bi-quartimin and bi-geomin rotations: First, the latter tends to outperform the former, particularly if cross-loadings or pure indicators are present. Second, both algorithms yield a severe overestimation of omega hierarchical values. As discussed previously, both methods tended to accommodate cross-loadings and other disturbances by moving variance from the group to the general factor (Mansolf & Reise, 2016), explaining the observed overestimation. As bi-quartimin is largely outperformed by bi-geomin due to the latter's ability to accommodate item complexity (as seen in Table 1 or in Abad et al., 2017, Table 2), it is of special interest to discuss this rotation method at greater length. Even though the geomin criterion has been praised in the literature (Mansolf & Reise, 2016), its application is far from simple, requiring researchers to be aware of its unique characteristics: (a) geomin (and bi-geomin by extension) is not a unique, single criterion, but a family of criteria dependent on the epsilon parameter (Hattori et al., 2017); (b) geomin often requires that researchers manually explore several local minima solutions to ensure that they find the simplest structure (Hattori et al., 2017; Mansolf & Reise, 2016); and c) under the bi-factor model, and when combined with the gradient projection algorithm, factor collapse occurrence has been routinely observed (Mansolf & Reise, 2016; Robertson, 2019), with limited information in the literature regarding under which conditions this is most likely to occur. Thus, practitioners should be aware of its limitations and provide sufficient justification for its application (i.e., how the epsilon parameter was decided, how local minima solutions were explored).

Limitations and future directions

The findings of this study only pertain to the conditions under study, and readers should proceed with caution when extending these results to alternative situations. In this study, only bi-factor structures with orthogonal factors were studied. The interpretation of oblique bi-factor structures is still controversial (Reise et al., 2018), and therefore remains unexplored. Furthermore, the effect of having factors varying in the number of indicators within the same structure was not considered here but is of interest in applied settings.

A significant limitation is that this article did not present evidence relative to how BEFA algorithms would perform with regards to the estimation of the reliability of scores from group factors (i.e., omega hierarchical subscale). The reason for this decision is that, to the best of our knowledge, no procedure for estimating the reliability of such scores in exploratory settings has yet been developed. Even though the omega hierarchical subscale has been successfully examined in the context of confirmatory models or the SL solution (Rodriguez et al., 2016b), translating a similar evaluation to assess the studied structures would involve a decision on which loadings are to be taken into account when computing each group factor variance. Therefore, it would involve either some kind of quasi-confirmatory approach within BEFA or applying cutoff points to define loading significance (which is known to be a questionable decision; Garcia-Garzon et al., 2019). Readers engaged in this approach might benefit from exploring these quasi-confirmatory models. The most common approach to these quasi-confirmatory models is to assign each item to the specific factor on which they load the highest and compute each omega hierarchical subscale using this solution. This approach is applied by the functions omegaSem and omega in the *psych* package (Revelle, 2018). As noted by a reviewer, omega hierarchical values would tend to be higher for these solutions than for the exploratory counterparts. Nevertheless, the relationship between omega hierarchical estimated from fully or semi-exploratory approaches should be investigated in future research.

Lastly, researchers have often been interested in additional aspects of the BEFA model apart from general factor score reliability. Among them, the main preoccupation is to evaluate the departure of such a model from a unidimensional model. The primary statistic applied in this context is the explained common variance (i.e., ECV; Rodriguez et al., 2016b), or the ratio between variance due to the general factor and the total common variance. This statistic was not evaluated due to two considerations: (a) ECV limitations as an indicator of the presence of a general factor (Revelle & Wilt, 2013); and (b) its relationship with omega hierarchical has been previously demonstrated to be mediated by the percentage of uncontaminated correlations (PUC; Reise et al., 2013). Unfortunately, and similarly to the case regarding omega hierarchical subscale, PUC has not yet been extended to BEFA models. Therefore, future research to advance our understanding in these areas is encouraged.

In conclusion, this article has shown how the selection of an appropriate BEFA algorithm could have a severe impact on a researcher's view on assessing whether a general factor sufficiently explains the test variance. This information could be crucial for researchers to evaluate claims regarding how reliable general factors genuinely are in a given field or test, especially if models are fitted using different algorithms and techniques. If researchers are interested in investigating BEFA models themselves, we recommend that applied researchers consider the application of the SLiD algorithm, as it is the most reliable of the techniques we examined over the broadest range of conditions and structures.

Article information

Conflict of interest disclosures: Each author signed a form for disclosure of potential conflicts of interest. No authors reported any financial or other conflicts of interest in relation to the work described.

Ethical principles: The authors affirm having followed professional ethical guidelines in preparing this work. These guidelines include obtaining informed consent from human participants, maintaining ethical treatment and respect for the rights of human or animal participants, and ensuring the privacy of participants and their data, such as ensuring that individual participants cannot be identified in reported results or from publicly available original or archival data.

Role of the funders/sponsors: None of the funders or sponsors of this research had any role in the design and conduct of the study; collection, management, analysis, and interpretation of data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Funding: This work was supported by Grant FPU15/03246 from the Ministry of Education, Culture and Sports (Spain), by Grant PSI2013-44300-P from Ministerio de Economia y Competitividad (Spain) and Fondo Social Europeo and by Grant PSI2017-85022-P from Ministerio de Economia y Competitividad (Spain) and Fondo Social Europeo.

Acknowledgements: The ideas and opinions expressed herein are those of the authors alone, and endorsement by the authors' institutions or the funding agencies is not intended and should not be inferred.

ORCID

Luis E. Garrido http://orcid.org/0000-0001-8932-6063

References

Abad, F. J., Garcia-Garzon, E., Garrido, L. E., & Barrada, J. R. (2017). Iteration of partially specified target matrices: Application to the bi-factor case. *Multivariate Behavioral Research*, *52*(4), 416–429. doi:10.1080/00273171.2017.1301244

Bechtoldt, H. P. (1961). An empirical study of the factor analysis stability hypothesis. *Psychometrika*, *26*(4), 405–432. doi:10.1007/BF02289771

Bentler, P. M. (2009). Alpha, dimension-free, and model-based internal consistency reliability. *Psychometrika*, *74*(1), 137–143. doi:10.1007/s11336-008-9100-1

Bernaards, C., & Jennrich, R. I. (2005). Gradient projection algorithms and software for arbitrary rotation criteria in factor analysis. *Educational and Psychological Measurement*, *65*(5), 676–696. doi:10.1177/0013164404272507

Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research*, *36*(1), 111–150. doi:10.1207/S15327906MBR3601_05

Caspi, A., & Moffitt, T. E. (2018). All for one and one for all: Mental disorders in one dimension. *American Journal of Psychiatry*, *175*(9), 831–844. doi:10.1176/appi.ajp.2018.17121383

Chen, F. F., Hayes, A., Carver, C. S., Laurenceau, J.-P., & Zhang, Z. (2012). Modeling general and specific variance in multifaceted constructs: A comparison of the bifactor model to other approaches. *Journal of Personality*, *80*(1), 219–251. doi:10.1111/j.1467-6494.2011.00739.x

Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bifactor and second-order models of quality of life. *Multivariate Behavioral Research*, *41*(2), 189–163. doi:10.1207/s15327906mbr4102_5

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Lawrence Erlbaum

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334. doi:10.1007/BF02310555

Garcia-Garzon, E., Abad, F. J., & Garrido, L. E. (2019). Improving bi-factor exploratory modelling: Empirical target rotation based on loading differences. *Methodology*, *15*(2), 45–55. doi:10.1027/1614-2241/a000163

Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F., Hothorn, T. (2017). *mvtnorm: Multivariate normal and t distributions.* R package version 1.08. https://CRAN.R-project.org/package=mvtnorm.

Giordano, C., & Waller, N. G. (2019). Recovering bifactor models: A comparison of seven methods. *Psychological Methods*, doi:10.1037/met0000227

Hancock, G. R., & Mueller, R. O. (2011). The reliability paradox in assessing structural relations within covariance structure models. *Educational and Psychological Measurement*, *71*(2), 306–324. doi:10.1177/0013164410384856

Harman, H. H. (1967). *Modern factor analysis* (2nd ed.). University of Chicago Press.

Hattori, M., Zhang, G., & Preacher, K. J. (2017). Multiple local solutions and geomin rotation. *Multivariate Behavioral Research*, *52*(6), 720–731. doi:10.1080/00273171.2017.1361312

Holzinger, K. J., & Swineford, F. (1937). The bi-factor method. *Psychometrika*, *2*(1), 41–54. doi:10.1007/BF02287965

Jamovi Project (2019). jamovi (Version 1.0.0.0) [Computer Software]. https://www.jamovi.org

Jennrich, R. I., & Bentler, P. M. (2011). Exploratory bi-factor analysis. *Psychometrika*, *76*(4), 537–549. doi:10.1007/s11336-011-9218-4

Jennrich, R. I., & Bentler, P. M. (2012). Exploratory bi-factor analysis: The oblique case. *Psychometrika*, *77*(3), 442–454. doi:10.1007/s11336-012-9269-1

Jensen, A. R., & Weng, L. J. (1994). What is a good g?. *Intelligence*, *18* (3), 231–258.(94)90029-9 doi:10.1016/0160-2896(94)90029-9

Mansolf, M., & Reise, S. P. (2016). Exploratory bifactor analysis: The Schmid-Leiman orthogonalization and Jennrich-Bentler analytic rotations. *Multivariate Behavioral Research*, *51*(5), 698–717. doi:10.1080/00273171.2016.1215898

McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, *23*(3), 412–433. doi:10.1037/met0000144

Moore, T. M., Reise, S. P., Depaoli, S., & Haviland, M. G. (2015). Iteration of partially specified target matrices: Applications in exploratory and Bayesian confirmatory factor analysis. *Multivariate Behavioral Research*, *50*(2), 149–161. doi:10.1080/00273171.2014.973990

Morin, A. J. S., Arens, A. K., & Marsh, H. W. (2016). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(1), 116–139. doi:10.1080/10705511.2014.961800

R Core Team (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

Raykov, T., & Marcoulides, G. A. (2019). Thanks coefficient alpha, we still need you!. *Educational and Psychological Measurement*, *79*(1), 200–210. doi:10.1177/0013164417725127

Reise, S. P., Bonifay, W., & Haviland, M. G. (2018). Evaluating scales bifactor modelling and the evaluation of scale scores. In P. Irwing, T. Booth, & D. J. Hughes (Eds.), *The Wiley handbook of psychometric testing: A multidisciplinary reference on survey, scale and test development* (1st ed., pp. 677–708). John Wiley & Sons, Ltd

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment*, *92*(6), 544–559. doi:10.1080/00223891.2010.496477

Reise, S. P., Morizot, J., & Hays, R. D. (2007). The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Quality of Life Research*, *16*(S1), 19–31. doi:10.1007/s11136-007-9183-7

Reise, S. P., Scheines, R., Widaman, K. F., & Haviland, M. G. (2013). Multidimensionality and structural coefficient bias in structural equation modeling: A bifactor perspective. *Educational and Psychological Measurement*, *73* (1), 5–26. doi:10.1177/0013164412449831

Revelle, W. (2018). *Psych: Procedures for personality and psychological research*. Northwestern University. https://CRAN.R-project.org/package=psch. Version 1.8.12.

Revelle, W., & Condon, D. M. (2018). Reliability. In P. Irwing, T. Booth, & D. J. Hughes (Eds.), *The Wiley handbook of psychometric testing: A multidisciplinary reference on survey, scale and test development* (1st ed., pp. 709–749). John Wiley & Sons

Revelle, W., & Condon, D. M. (2019). Reliability from α to ω: A tutorial. *Psychological Assessment*, *31*(12), 1395–1411. doi:10.1037/pas0000754

Revelle, W., & Wilt, J. (2013). The general factor of personality: A general critique. *Journal of Research in Personality*, *47*(5), 493–504. doi:10.1016/j.jrp.2013.04.012

Revelle, W., & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega, and the glb: Comments on sijtsma. *Psychometrika*, *74*(1), 145–154. doi:10.1007/s11336-008-9102-z

Robertson, S. (2019). *Bi-factor models and factor collapse: A Monte Carlo study* (Doctoral dissertation). Retrieved from Clemson University Institutional Repository.

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016a). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment*, *98*(3), 223–237. doi:10.1080/00223891.2015.1089249

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016b). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, *21*(2), 137–150. doi:10.1037/met0000045

Savalei, V., & Reise, S. P. (2019). Don't forget the model in your model-based reliability coefficients: A reply to McNeish (2018.). *Collabra: Psychology*, *5*(1), 36. doi:doi:10.1525/collabra.247

Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika*, *22*(1), 53–61. doi:10.1007/BF02289209

Schmitt, T. A., & Sass, D. A. (2011). Rotation criteria and hypothesis testing for exploratory factor analysis: Implicatons for factor pattern loadings and interfactor correlations. *Educational and Psychological Measurement*, *71*(1), 95–113. doi:10.1177/0013164410387348

Schönemann, P. H. (1966). A generalized solution of the orthogonal Procrustes problem. *Psychometrika*, *31*(1), 1–10. doi:10.1007/BF02289451

Thurstone, L. L. (1933). *The theory of multiple factors*. Edwards Brothers.

Viladrich, C., Angulo-Brunet, A., & Doval, E. (2017). A journey around alpha and omega to estimate internal consistency reliability. *Annals of Psychology*, *33*(3), 755–782. doi:10.6018/analesps.33.3.268401

Waller, N. G. (2018). Direct Schmid–Leiman transformations and rank-deficient loadings matrices. *Psychometrika*, *83*(4), 858–870. doi:10.1007/s11336-017-9599-0

Waller, N. G. (2019). *fungible: Psychometric functions from the Waller lab*. URL: http://doi.org/CRAN.R-project.org/package=fungible.Version1.86.

Yang, Y., & Green, S. B. (2015). Evaluation of structural equation modelling estimates of reliability for scales with ordered categorical items. *Methodology*, *11*(1), 23–34. doi:10.1027/1614-2241/a000087

Zinbarg, R. E., & Alden, A. R. (2015). Coefficient alpha and coefficient omega hierarchical. In Robin L. Cautin & Scott O. Lilienfeld (Eds.), *The Encyclopedia of Clinical Psychology* (1st ed., pp. 1–6). John Wiley & Sons, Ltd. doi:10.1002/9781118625392.wbecp543

Zinbarg, R. E., Revelle, W., & Yovel, I. (2007). Estimating ωh for structures containing two group factors: Perils and prospects. *Applied Psychological Measurement*, *31*(2), 135–157. doi:10.1177/0146621606291558

Zinbarg, R. E., Yovel, I., Revelle, W., & McDonald, R. P. (2006). Estimating generalizability to a latent variable common to all of a scale's indicators: A comparison of estimators for ωh. *Applied Psychological Measurement*, *30*(2), 121–144. doi:10.1177/0146621605278814

Appendix 1

Table A1. SLiD algorithm first iteration target matrix definition for Thurstone 9 mental tests.

Panel A: Original SL solution					Panel B: Squared normalized loadings				Panel E: First iteration target matrix			
Item	GF	GRF.1	GRF.2	GRF.3	GF	GRF.1	GRF.2	GRF.3	GF	GRF.1	GRF.2	GRF.3
Sen.	.71	.57	−.02	.03	–	1.00	.00	.00	NA	NA	0	0
Voc.	.73	.55	.04	−.02	–	.99	.01	.00	NA	NA	0	0
Com.	.68	.52	.03	.00	–	1.00	.00	.00	NA	NA	0	0
Fir.	.65	.00	.56	.00	–	.00	1.00	.00	NA	0	NA	0
Wor.	.62	−.01	.49	.08	–	.00	.98	.02	NA	0	NA	0
Suf.	.56	.11	.41	−.06	–	.07	.91	.02	NA	0	NA	0
Ser.	.59	.02	−.01	.61	–	.00	.00	1.00	NA	0	0	NA
Ped.	.58	.23	−.03	.34	–	.32	.01	.68	NA	NA	0	NA
Gro.	.54	−.04	.13	.46	–	.01	.08	.92	NA	0	0	NA
Average		.22	.18	.16		.38	.33	.29				

Panel C: Sorted squared normalized loadings*						Panel D: One-lagged differences distribution*					
Item	GRF1	Item	GRF2	Item	GRF3	Item	GRF1	Item	GRF2	Item	GRF3
Sen.	1.00	Fir.	1.00	Ser.	1.00	Sen.	.00	Fir.	.02	Ser.	.08
Voc.	1.00	Wor.	.98	Gro.	.92	Voc.	.00	Wor.	.06	Gro.	.24
Com.	.99	Suf.	.91	Ped.	.68	Com.	.68	Suf.	**.83**	Ped.	**.65**
Ped.	.32	Gro.	.08	Wor.	.02	Ped.	**.25**	Gro.	.07	Wor.	.00
Suf.	.07	Voc.	.01	Suf.	.02	Suf.	.06	Voc.	.00	Suf.	.02
Gro.	.01	Ped.	.01	Sen.	.00	Gro.	.01	Ped.	.00	Sen.	.00
Ser.	.00	Com.	.00	Voc.	.00	Ser.	.00	Com.	.00	Voc.	.00
Wor.	.00	Sen.	.00	Fir.	.00	Wor.	.00	Sen.	.00	Fir.	.00
Fir.	.00	Ser.	.00	Sen.	.00	Fir.	–	Ser.	–	Sen.	–
Average							*.13*		*.12*		.13

Note: Sen. = Sentences. Voc. = Vocabulary. Com. = Sentence completion. Fir. = First letters. Wor. = Four letter words. Suf.= Suffixes. Ser. = Letter series. Ped. = Pedigrees. Gro.= Letter group. GF = General Factor. GRF = Group Factor. Panels divided as item order is changed between Panels B and C. Substantive loadings appear shadowed in strong gray and cross-loadings in light gray. First one-lagged difference above factor difference average is presented bolded and underlined. * GF not presented in the Panel.

Appendix 2

Table A2. Marginal omega hierarchical root squared mean error (RMSE) each method in bi-factor structures.

Variable/level	Bi-geomin	Bi-quartimin	SL	SLiD	DSL	DBF
SAMPLE						
150	.097	.117	.127	**.079**	.098	.100
500	.072	.097	.079	**.042**	.091	.093
1000	.061	.084	.062	**.032**	.089	.091
NUM.GRF						
3	.115	.130	.072	**.071**	.106	.109
6	.041	.066	.107	**.031**	.079	.080
VAR.GRF						
4	.083	.102	.091	**.061**	.099	.100
8	.072	.093	.088	**.041**	.081	.088
CROSS.GRF						
No	.056	.053	.087	**.045**	.081	.082
Yes	.100	.143	.092	**.056**	.104	.107
PURE.GF						
No	.099	.116	.075	**.057**	.102	.105
Yes	.057	.080	.104	**.045**	.084	.084
SIZE.GF						
Low	.108	.137	.105	**.078**	.119	.126
Medium	.081	.102	.096	**.051**	.052	.053
High	.044	.055	.067	**.024**	.108	.104
STRUCTURE						
IC	.070	.067	.074	**.050**	.086	.089
ICB	.128	.165	.075	**.063**	.118	.122
ICP	.041	**.038**	.099	.041	.077	.075
ICBP	.073	.121	.109	**.049**	.091	.092
AVERAGE	.078	.098	.089	**.051**	.093	.094

Note. SAMPLE = sample size; NUM.GRF = number of group factors; VAR.GRF = number of indicators per group factor; CROSS.GRF = Cross-loading presence; PURE.GF = Pure indicator presence; SIZE.GF = General factor average factor loading; IC = Independent Cluster; ICB = Independent Cluster Basis; ICP = Independent Cluster with Pure Indicators; ICBP = Independent Cluster Basis with Pure Indicators; SL = Schmid-Leiman orthogonalization; SLiD = Iterative Empirical Target Rotation based on an initial Schmid-Leiman solution; DSL = Direct Schmid-Leiman. DBF = Direct Bi-factor. MBE values appear under brackets. Conditions with MAE ≥ .05 appear shadowed in gray. Lowest MAE for each condition appears bolded.

Table A3. Marginal omega hierarchical root squared mean error (RMSE) across each method for second-order models.

Variable / Level	Bi-geomin	Bi-quartimin	SL	SLiD	DSL	DBF
SAMPLE						
150	.098	.114	.132	**.086**	.082	.085
500	.075	.072	.067	**.050**	.070	.073
1000	.074	.055	.044	**.043**	.068	.073
NUM.GRF						
3	.118	.104	**.065**	.080	.084	.087
6	.048	.056	.096	**.039**	.064	.065
VAR.GRF						
4	.082	.083	.082	**.069**	.072	.076
8	.084	.077	.079	**.050**	.075	.076
SIZE.SOF						
Low	.106	.100	**.075**	.082	.137	.147
Medium	.080	.079	.082	.055	**.016**	.022
High	.063	.062	.085	**.042**	.067	.059
AVERAGE	.83	.080	.081	**.060**	.074	.076

Note. SAMPLE = sample size; NUM.GRF = number of group factors; VAR.GRF = number of indicators per group factor; SIZE.SOF = Second-order general factor average factor loading; SL = Schmid-Leiman orthogonalization; SLiD = Iterative Empirical Target Rotation based on an initial Schmid-Leiman solution; DSL = Direct Schmid-Leiman. DBF = Direct Bi-factor. MBE values appear under brackets. Conditions with MAE ≥ .05 appear shadowed in gray. Lowest MAE for each condition appears bolded.

Table A4. Marginal omega hierarchical root mean squared error (RMSE) across methods for factor structures without a general factor.

Variable/Level	Bi-geomin	Bi-quartimin	SL	SLiD	DSL	DBF
SAMPLE						
150	.277	.271	.177	**.133**	.501	.529
500	.264	.210	.160	**.088**	.497	.513
1000	.261	.195	.152	**.077**	.496	.509
NUM.GRF						
3	.360	.288	.213	**.142**	.463	.487
6	.175	.162	.114	**.057**	.534	.547
VAR.GRF						
4	.242	.233	.160	**.122**	.447	.474
8	.293	.218	.167	**.077**	.549	.561
CROSS.GRF						
No	.250	.124	.134	**.078**	.490	.511
Yes	.263	.326	.192	**.120**	.506	.524
AVERAGE	.267	.225	.163	**.099**	.498	.517

Note. SAMPLE = sample size; NUM.GRF = number of group factors; VAR.GRF = number of indicators per group factor; CROSS.GRF = Cross-loading presence; PURE.GF = Pure indicator presence; IC = Independent Cluster; ICB = Independent Cluster Basis; ICP = Independent Cluster with Pure Indicators; ICBP = Independent Cluster Basis with Pure Indicators; SL = Schmid-Leiman orthogonalization; SLiD = Iterative Empirical Target Rotation based on an initial Schmid-Leiman solution; DSL = Direct Schmid-Leiman. DBF = Direct Bi-factor. MBE values appear under brackets. Conditions with MAE ≥ .05 appear shadowed in gray. Lowest MAE for each condition appears bolded.

Chapter 6

Bi-factor Exploratory Structural Equation Modelling Done Right

This study is currently submitted to *Psicothema*. The following pages include the latest version of this article. This article was co-authored with Nieto, M.D. (Universidad de Nebrija, Spain), Garrido, L.E. (Pontificia Universidad Católica Madre y Maestra, Dominican Republic) and Abad, F.J. (Universidad Autónoma de Madrid, Spain).

Abstract

Background: Due to its flexibility and superior statistical properties, bi-factor Exploratory Structural Equation Modeling (ESEM) has become a preferred tool in psychometrics. Unfortunately, most recent methods for approximating these structures, such as the SLiD algorithm, are not available in the principal software for conducting ESEM (i.e., *Mplus*). To resolve this issue, a novel, user-friendly Shiny application for integrating the SLiD algorithm in bi-factor ESEM estimation in *Mplus* is presented. Thus, a two-stage framework for conducting SLiD-based bi-factor ESEM in *Mplus* was developed. **Method:** This approach is presented in a step-by-step guide for applied researchers, showing the utility of the developed SLiDApp application. Using data from the Open-Source Psychometrics Project (N = 2495), we conduct a bi-factor ESEM exploration of the Generic Conspiracist Beliefs Scale. We studied whether bi-factor modelling was appropriate and if both, general and group factors, were related to each personality traits. **Results:** It was further exemplified how the SLiD algorithm provided unique information regarding its factor structure and structural parameters. **Conclusions:** The results illustrated the usefulness and validity of SLiD-based bi-factor ESEM, and how the proposed Shiny app could facilitate the use of these methods for applied researchers.

Keywords: bi-factor, exploratory structural equation modelling, factor analysis; rotation, Schmid-Leiman.

Resumen

Antecedentes: Debido a sus propiedades estadísticas, los modelos bi-factoriales de ecuaciones estructurales exploratorias (bi-factor ESEM) son una herramienta clave en psicometría. Desafortunadamente, las últimas alternativas para su estimación no se encuentran disponibles en el software principal usado para su estimación (i.e., *Mplus*). Para solucionar este problema, se presenta una aplicación Shiny (SLiDApp) que permite integrar los resultados del algoritmo SLiD en un modelo bi-fctor ESEM estimado en *Mplus*. Así, se diseñó una estrategia de dos pasos para estimar modelos bi-factor ESEM basados en SLiD. **Método:** Este enfoque se ilustra mediante una guía paso por paso que muestra cómo usar la aplicación diseñada para este fin. Se realizó un modelo bi-factor ESEM basado en SLiD de la Escala de Creencias Conspirativas Genéricas usando una muestra del Open-Source Psychometrics Project (N = 2495). Se analizó la relación de los factores generales y de grupo con los cinco factores de personalidad. **Resultados:**. Los resultados mostraron cómo el algoritmo SLiD proveía de información única acerca de la estructura factorial y los parámetros estructurales. **Conclusiones:** Este estudio demostró la utilidad tanto de los modelos bi-factoriales ESEM basados en SLiD cómo de la app propuesta. Asimismo se espera que esta aplicación facilite el uso de este tipo de métodos por parte de investigadores aplicados.

Palabras clave: bi-factor, modelos de ecuaciones estructurales exploratorias, análisis factorial; rotación, Schmid-Leiman.

The bi-factor model plays today a crucial role in the advancement of psychological theory (Reise, Bonifay, Haviland, 2017) with major applications in personality and intelligence (Garcia-Garzon, Abad, Garrido, 2019a; Primi, da Silva, Rodrigues, Muniz, Almeida, 2013). Bi-factor models represent a convenient set of factor models that allows the simultaneous estimation of a general factor (common to all items) alongside with several group factors (underlying to specific sets of items; Reise, Bonifay, Haviland, 2017). As such, bi-factor models have been recently introduced in the context of Exploratory Structural Equation Modeling (i.e., ESEM; Gomes, Almeida, Núñez, 2017). ESEM has recently gained popularity as it has been shown to improve parameter estimation when compared with traditional structural equation modelling (Guo et al., 2019; Marsh, Guo, Dicke, Parker, Craven, 2019).

The principal ESEM feature is the introduction of Exploratory Factor Analysis (i.e., EFA) measurement models within a SEM model while retaining global and local fit inspection and the ability to include residual correlations in the measurement model (Nieto et al., 2017; Asparouhov Muthén, 2009; Garrido et al., 2018). Accordingly, a decisive step in ESEM is, as in any EFA-based model, the choice of an appropriate rotation method (Izquierdo, Olea, Abad, 2014). Such a decision might be of more relevance in this context, as any estimation bias present in the measurement model propagates to other parameters in the model (Guo et al., 2019; Reise et al., 2017).

With regards to bi-factor modelling, several rotation alternatives are currently available (Abad, Garcia-Garzon, Garrido, Barrada, 2017; Asparouhov Muthén, 2009; Garcia-Garzon, Abad, Garrido, 2019b; Giordano Waller, 2019; Lorenzo-Seva Ferrando, 2018). In this sense, this article is designed to introduce the use of one of the current state-of-the-art bi-factor rotation methods within ESEM: the Empirical Iterative Target Rotation based on a Schmid-Leiman solution (Garcia-Garzon et al., 2019b). As this method is only available in R software and ESEM is primarily conducted using *Mplus* (Muthén Muthén, 2017), a novel friendly-user Shiny application was develop to integrate both softwares (called SLiDApp). Its utility is illustrated by a step-by-step guide and an empirical bi-factor ESEM analysis of the Generic Conspiracy Belief Scale (GCBS; Brotherton et al., 2013).

The SLiD Algorithm

As interest in bi-factor exploratory factor analysis (i.e., BEFA) has dramatically grown over the last decade, many articles have been concerned with studying their application within ESEM (Asparouhov Muthén, 2009). The principal software to conduct ESEM is *Mplus* (Muthén Muthén, 2017), which offers three approaches towards estimating BEFA models in this context: bi-quartimin, bi-geomin (Jennrich Bentler, 2011, 2012) and the non-iterative target rotation (Reise, Moore, Maydeu-Olivares, 2011). Unfortunately, it is well known in the BEFA literature that these approaches present stringent limitations and fail to provide accurate parameter estimation under most realistic conditions (Abad et al., 2017; Garcia-Garzon, Abad Garrido, in press; Giordano Waller, 2019).

Accordingly, several alternatives have recently appeared in the literature: the Direct Schmid-Leiman, the Direct Bi-factor (Giordano Waller, 2019), the Pure Exploratory Bi-factor Analysis (PEBI; Lorenzo-Seva Ferrando, 2018) and the Empirical Iterative Target Rotation based in a Schmid-Leiman Solution (i.e., SLiD; Garcia-Garzon, Abad, Garrido, 2019b).

Amongst those, the SLiD algorithm presents a unique combination of features (Garcia-Garzon, Abad, Garrido, 2019b). The SLiD algorithm has been shown to result in both, improved parameter estimation when compared with alternative algorithms (Garcia-Garzon, Abad Garrido, 2019b) and unbiased estimation of general factor reliability under many circumstances (Garcia-Garzon, Abad Garrido, in press). The SLiD algorithm approximates a simple exploratory bi-factor model in four main steps: (a) First, an initial Schmid-Leiman model is estimated, which is known to represent a biased estimation of the bi-factor model of interest (Reise, Moore Maydeu-Olivares, 2011); (b) Second, the initial Schmid-Leiman solution is used to define a partially specified target matrix using an empirical, factor-specific cut-off point based on loadings' differences (as detailed in Garcia-Garzon, Abad Garrido., 2019b); (c) An initial, tentative exploratory bi-factor solution is computed employing a target rotation using the empirically defined target matrix; (d) The estimated bi-factor solution is subsequently refined through repeating steps b and c until convergence (i.e., the target rotation becomes stationary within iterations); (e) Finally, the refined structure is modified so to approximate the identification conditions defined in Asparouhov and Muthén (2009).

An additional benefit of the SLiD algorithm is that it is freely available in open-source software such as R, which facilitates its integration into alternative platforms and applications. Unfortunately, as said before, the SLiD algorithm is not available in *Mplus* (Muthén Muthén, 2017), which is the preferred software to conduct ESEM. Thus, as of today, practitioners wishing to apply a bi-factor ESEM face an uncomfortable situation: (a) to conduct this analysis using a detrimental rotation method such as bi-geomin (the default option in *Mplus*), which would lead to biased, incorrect results; (b) to pre-estimate the measurement models using R using the SLiD algorithm and to translate the rotated factor solutions as fixed parameters in a traditional structural equation model; (c) a two-step framework for computing state-of-the-art bi-factor ESEM, where a refined target bi-factor rotation matrix is estimated in R using the SLiD algorithm and is subsequently used in *Mplus* to perform ESEM (as in Garcia-Garzon et al., 2019a). Unfortunately, researchers interested in this latter option would be required to be familiarized with both, R and *Mplus* software. Thus, to bridge the gap between both software, and provide users with an easy pathway to apply this two-step framework to perform SLiD-based ESEM, a novel Shiny app was developed.

SLiDApp: Implementing Modern BEFA in ESEM Models

As previously acknowledged in this journal, methodological innovations such as the SLiD algorithm are only useful to the extent that they are implemented in software available to the general public (Calderón-Garrido, Navarro-González, Lorenzo-Seva, Ferrando-Piera, 2019). To this end, in recent years Shiny-based web applications are gaining popularity (e.g., Nieto, Garrido, Golino, Shi, Abad, 2019). Shiny is an R package that allows developing interactive web tools (Chan, Cheng, Allaire, Xie, McPherson, 2019). This article will introduce the *SLiDapp* (https://slidapp.shinyapps.io/SLid_app/), a user-friendly Shiny application that provide the refined bi-factor target resulting from the SLiD algorithm in a format ready to be introduced in *Mplus* and applied within an ESEM context (Figure 1).

- PLEASE INSERT FIGURE 1 HERE –

The different steps to use the app and its features are illustrated in Figure 1. These steps are further illustrated in the Instructions panel within the application. The first step is to select a file in TXT or DAT format including variables to be analyzed. The next steps are concerned with file characteristics, such as whether variables names are included in the header (step 2) or the separator character applied (step 1). If the data-set contains missing values, the user must specify how they are coded in the input box shown in step 4 (multiple missing values are accepted). Afterwards, step 5 consists of loading the dataset to the SLiDapp using the "Load Data" option. The user can preview the loaded data using the Data Preview box (step 6) or by clicking in Display data (step 7).

To start the analysis, the researcher must specify the number of group factors to be extracted (step 8). In this case, a SLiD solution requires at least two group factors to be estimated. After deciding on the model dimensionality, the "Run SLid" option will be now clickable to run the SLiD algorithm (step 9). A progress bar will be shown while SLiD finishes the computation of the target matrix (step 10). Finally, the estimated solution will be printed on the app interface, and ready to be copied (step 11) and/or inserted in a *Mplus* input file. Interested users can save the estimated target matrix in their computers by indicating a name for the resulting file (step 12) and clicking on the "Download" option (step 13). Furthermore, users can download the SLiD output in CSV format if interested (step 14). The output produced by the app contains the ANALYSIS and MODEL sections of a *Mplus* input file. Furthermore, it controls that each line does not exceed 90 characters (a *Mplus* restriction). Thus, users only need to add the appropriate code regarding the structural part of the estimated model and to adapt the code of the remaining sections (i.e., DATA, VARIABLE and OUTPUT) in the input and output file (see Supplementary data for a reproducible example using the GBCS data).

The SLiD algorithm is run automatically using either polychoric or Pearson's correlation based on the number of categories detected in the variables, applying the unweighted least-squares extraction method and the oblimin rotation when estimating the initial Schmid-Leiman solution. The utility of this Shiny app is illustrated below by conducting a SLiD based bi-factor ESEM to the Generic Conspiracist Belief Scale (GCBS; Brotherton, French, Pickering, 2013). This example would investigate both, whether a bi-factor model holds for the GCBS and the relationship between the general and group factors and personality traits derived from the big five models (Goreis Voracek, 2019). To illustrate the complete process of computing a bi-factor ESEM model, dimensionality assessment and factor structure choice are also presented below.

The utility of this Shiny app is illustrated below by conducting a SLiD based bi-factor ESEM to the Generic Conspiracy Belief Scale (GCBS; Brotherton, French, Pickering, 2013). This example would investigate both, whether a bi-factor model holds for the GCBS and the relationship between the general and group factors and personality traits derived from the big five model (Goreis Voracek, 2019). To illustrate the complete process of computing a bi-factor ESEM model, dimensionality assessment and factor structure choice are also presented below.

Method

The GCBS (Brotherton et al., 2013) represents the primary assessment tool in research areas such as inquiring beliefs in fake news, beliefs in conspiracy theories and new forms of information consumption. Accordingly, it has more than 33 research applications in the last five years (Goreis Voracek, 2019; Hollander, 2018). In this area, there is an increasing controversy surrounding whether conspiracy beliefs are correlated with individual aspects such as personality traits (for a detailed review, see Goreis Voracek, 2019). While some authors have suggested that higher tendency to believe in conspiracies theories are linked with lower agreeableness and emotional stability and higher openness to experience (Brotherton et al., 2013; Goreis Voracek, 2019), other argued that such effects were equivocal, to say at least (Goreis Voracek, 2019).

There are many reasons to believe that the literature surrounding GCBS could benefit from an exploration of bi-factor ESEM. Firstly, even though the GCBS was developed as a multidimensional 15-item tool assessing five different conspiracy believes domains (Brotherton et al., 2013), it has been primarily applied as a unidimensional scale assessing a general, conspiracist ideation factor (Hollander, 2018; Swami et al., 2017). Despite the theoretical support for the idea of a general conspiracy ideation factor (Goertzel, 2013; Swami et al., 2017; Wood, Douglas, Sutton, 2012), evidence showed that unidimensional (or even two-dimensional) GCBS models presented substantive fit issues (Brotherton et al., 2013; Swami et al., 2017). Thus, the latent GCBS structure is still a matter of debate in the literature (Swami et al., 2017). In this sense, a bi-factor model could help to understand the extent that GCBS represents an essentially unidimensional tool and whether the group scales reflect any relevant information additional to this general factor (Reise, Bonifay, Haviland, 2017; Rodriguez, Reise, Haviland, 2016).

Participants

Using data from the Open-Source Psychometric Project (www.openpsychometrics.org), responses of 2495 individuals who responded online to the GCBS, the ten-item personality inventory (i.e., TIPI; Gosling, Rentfrow, Swann, 2003) and several demographic items were analyzed. The sample was gender-balanced (females represented 49.0% of the sample), consisted of young aged (M = 27.63, SD = 13.36), higher-educated (36.9% completed university-level studies), English-native speakers (75.2% of participants). Thirteen respondents were removed from the sample due to having response times over 30 minutes (response times over two minutes per item).

Instruments

The GCBS is a short, 15-item scale that assesses five generic conspiracy domains: government malfeasance, extraterrestrial cover-up, malevolent global conspiracies, personal well-being, and control of information. All items are measured on a five-point Likert Scale to evaluate the veracity of given sentences (i.e., “Evidence of alien contact is being concealed from the public”) ranging from 1 (“definitely not true”) to 5 (“definitely true”). The complete item descriptions are offered in the original manuscript (Table A1; Brotherton et al., 2013).

The TIPI is a brief personality measure which assesses the Big Five personality model (including extraversion, openness to experience, agreeableness, conscientiousness, and emotional stability traits), asking individuals to rate themselves with regards to five positive and five negative adjectives applying a Likert scale ranging from 1 (“Disagree strongly”) to 7 (“Agree strongly”). After reversing responses to negative items, each personality trait is measured as the average of the two corresponding items.

Data Analysis

A complete factor-analysis study was conducted to illustrate all the necessary steps appropriate for conducting bi-factor ESEM. Firstly, GCBS dimensionality was estimated employing parallel analysis (Garrido, Abad, Ponsoda, 2013). Afterwards, unidimensional, confirmatory and exploratory versions of the five correlated factors and confirmatory and exploratory bi-factor models with five group factors were analyzed. Secondly, several bi-factor rotation methods were tested, namely the bi-geomin, bi-quartimin, a theory-driven partially specified target rotation, and a SLiD-based target rotation. The quality and reliability of each solution were assessed through omega hierarchical (i.e., H), the expected common variance (i.e., ECV), and the replicability index (i.e., H-index) following Rodriguez, Reise, and Haviland (2016) guidelines. Lastly, a bi-factor ESEM model using the SLiD-based target rotation was conducted to estimate the relationship between the different GCBS factors and TIPI personality traits. In these analyses, the SLiD-based target was estimated using the Shiny app. All subsequent analyses were performed using the weighted least-squares with mean and variance correction (WLSMV) in *Mplus* 7 (Muthén Muthén, 2007). Parallel analysis and bi-factor indices were computed in R 3.6.2 (R Core Team, 2019) using the psych package 1.9.12. (Revelle, 2019). Due to data characteristics (i.e., few items per factor and expected high inter-factor correlations) analysis was conducted over the reduced polychoric correlation matrix using the mean eigenvalue rule to decide the number of appropriate factors (Golino et al., 2020).

Results

GCBS Exploratory Factor Analysis

Even though parallel analysis indicated that five factors should be retained (empirical eigenvalues were 8.30, .83, .40, .14, .08 and .02, and average resampled eigenvalues were .36, .12, .11, .08, .06 and .05), the relative size of the first eigenvalue indicates that a dominant dimension might be present. Thus, this dimensionality assessment suggested a combination of a strong single factor altogether with additional minor factors consistent with the hypothesis of a bi-factor model being appropriate for GCBS.

Five different GCBS measurement models were compared in terms of data fit (Table 1), from which an EFA model with 5 factors fitted the data the best. RMSEA was observed to notably differ between CFA and EFA models, which could be attributed to differences of magnitude in the inter-factor correlation matrices. Under these circumstances, the latter was favoured. An exploratory version of the latter model showed an adequate fit to the data

(CFI > .99; TLI > .99; RMSEA < .05), but presented a complex factor pattern (i.e., with 7 out of 15 items presenting cross-loadings larger than .20 in absolute value). The high inter-factor correlations, with 5 out of 10 inter-factor correlations having values over .60, indicate substantive overlap over the five correlated factors. Thus, it was decided to explore the fit of bi-factor models combining a general factor plus five additional group factors.

- PLEASE INSERT TABLE 1 HERE –

A Comparison of BEFA Algorithms

Several BEFA solutions were explored, namely a bi-quartimin, bi-geomin, a model rotated using partial target rotation towards the theoretical structure, and a SLiD-based rotation solution (Table 2). Substantive differences across models were found with regards to group factor loadings pattern: Firstly, from *Mplus* rotation criteria, bi-geomin produced the simplest solution. Unfortunately, items 9, 5 and 14 only presented substantive loadings (larger than |.20|) in the general factor in this solution, leaving the personal well-being factor defined by a single item (Item 4). Secondly, the theoretical target did not perform any better than this model. Thirdly, the only solution properly recovering the personal well-being factor was SLiD (see Figure 1 for the specified target matrix), with items 9 and 14 substantially loading on it. As SLiD was the structure with a stronger resemblance to the theoretical expectation, it was subsequently retained as the model to be used in ESEM analyses. Noteworthy, all models supported the presence of a reliable general factor (ω_H, ECV and H-index: bi-quartimin .91/.74/.95; bi-geomin = .92/.76/.95; theory-based target = .91/.75/.95; SLiD = .92./.76/.95).

- PLEASE INSERT TABLE 2 HERE –

The SLiD-Based Bi-factor ESEM Model

After deciding on a bi-factor measurement model, the SLiD-based target matrix was applied within an ESEM framework in *Mplus* via to the Shiny app. In this model, we explored the relationships between the GCBS factors and the Big Five personality traits utilizing ESEM (Figure 2).

- PLEASE INSERT FIGURE 2 HERE –

The standardized regression parameter and the explained variance (R^2) for each personality factor are presented in Table 3. The model fitted the data excellently (CFA = 1.000; TLI > .99; RMSEA < .02; Table 2). This ESEM model revealed distinct patterns of relationships for each GCBS factor involved, such as the general factor being the only factor significantly (and positively) related with openness (β = .105, p < .001) and negatively related with both agreeableness (β = -.058, p < .001) and emotional stability (β = -.076, p < .001). Nevertheless, the observed relationships presented small predictive power (.013 < R^2 < .038). Thus, even though the results from the bi-factor ESEM results aligned with some theoretical predictions, they largely supported previous conclusions suggesting that the relationships between the GCBS factors and personality traits are, at best, weak (Goreis Voracek, 2019).

- PLEASE INSERT TABLE 3 HERE –

Finally, it should be highlighted that whether another bi-factor rotation had been chosen, these results would have been substantially different. Table A1 (Appendix 1) shows the regression parameters from ESEM models obtained with bi-quartimin, bi-geomin, and the theory-based target rotation. Notable disagreements between methods were found among methods, particularly, but not limited, to the personal well-being and control of information group factors. For example, the relationship between conscientiousness and personal well-being was observed to be significantly positive bi-quartimin, non-significant for bi-geomin and the theory target and significantly positive for and the SLiD-based target. In the case of control of information, SLiD provided was the only method to not found any significant relationship between personality traits and GCBS factors. Thus, these results reflect that the bi-factor rotation critically determined the nature and direction of the estimated structural parameters.

Discussion

Bi-factor ESEM models constitute today a decisive tool for latent variable modelling. As such, many researchers who have become interested in these models in the literature find themselves limited to certain bi-factor rotation methods as these models are primarily estimated using *Mplus*. Despite their popularity, software default methods are not always appropriate, and ultimately impair the advancement of good analytical practices in the context of factor analysis (Izquierdo et al., 2014).

To improve this situation, this research aimed to provide readers with the necessary tools for applying modern bi-factor estimation utilizing the SLiD algorithm within ESEM. We exemplified how to perform this analysis by illustrating the use of a Shiny application and a novel bi-factor ESEM exploration of the relationship between the Generic Conspiracist Beliefs Scale and the big five personality traits. Results evidenced the relationships between GCBS factors and personality traits were dependent upon the choice of the bi-factor rotation methods. Moreover, despite supporting latest findings in the GCBS literature (Goreis Voracek, 2019), it is our understanding that such scale would ultimately benefit from being re-constructed following current directions in the field (Muñiz Fonseca-Pedrero, 2019).

Lastly, this study was not without limitations. For example, as of today, the Shiny app does not allow users to choose an estimation method or initial rotation method for SLiD. However, future versions of the app will expand these capabilities. Moreover, it should be acknowledged that other relevant bi-factor rotation methods, such as the Pure Exploratory Bi-factor Analyses, were not explored here (as they are only available in specialized software; FACTOR; Ferrando Lorenzo-Seva, 2017; Lorenzo-Seva Ferrando, 2018).

With bi-factor ESEM being posed to play a substantial role in future psychological research, we consider of importance to ensure that interested researchers can use state-of-the-art methods regardless of their programming skills. With the same humble spirit that other colleagues have previously expressed in this journal (Calderón Garrido et al., 2019), it could only be hoped that by facilitating the use of these modern methods via a user-friendly, free Shiny app, we have contributed to foster critical thinking and good researcher practices in the context of factor analysis.

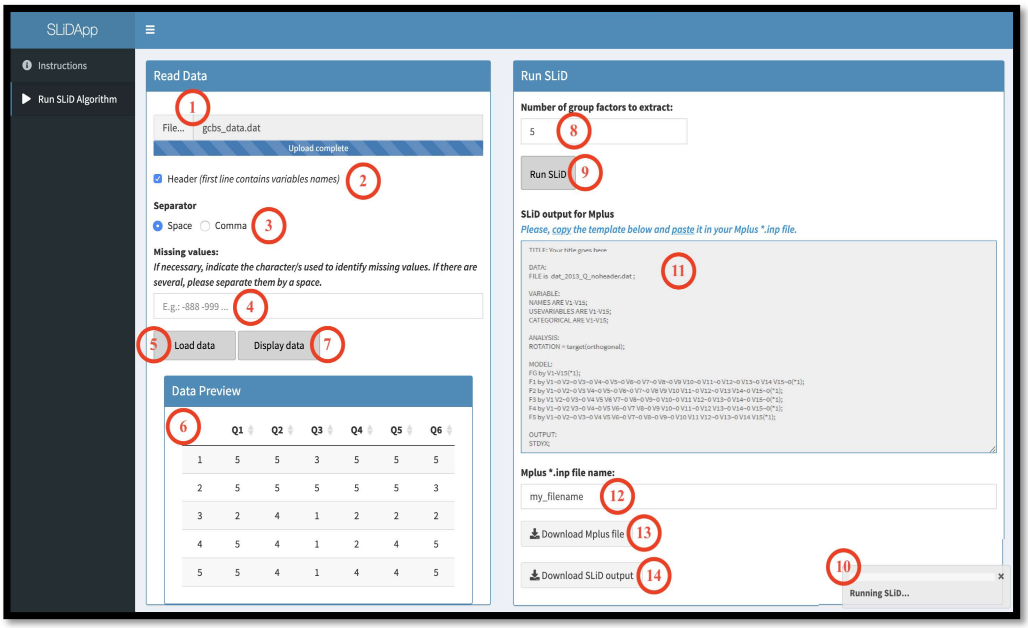

Figure 1. The interface of the Shiny application for computing a SLiD-based target matrix. The different steps for using the app are highlighted and circled in red.

Table 1.

Model fit indices for confirmatory and exploratory models estimated.

	χ^2	*df*	*p*	CFI	TLI	RMSEA
Unidim.	6760.73	90	.000	.922	.909	.172 (.169 - .176)
CFA 5 factors correlated	1018.30	80	.000	.989	.986	.069 (.065 - .072)
EFA 5 factors correlated	192.56	40	.000	.998	.995	.039 (.034 - .045)
CFA bi-factor	1158.64	75	.000	.987	.982	.076 (.072 - .080)
EFA bi-factor	61.67	30	.001	1.000	.999	.021 (.013 - .028)
ESEM	*115.85*	*75*	*.002*	*1.000*	*.999*	*.015 (.009 - .020)*

Note. Unidim. = Unidimensional model; X^2= Chi-square statistic; *df* = degrees of freedom; *p*= p-value associated with X^2 test of fit; CFI = Comparative fit index; TLI= Tucker-Lewis index; RMSEA = Root Mean Square Error of Approximation (with 95% confidence interval in parenthesis); CFA = Confirmatory Factor Analysis; EFA = Exploratory Factor Analysis; ESEM = Exploratory Structural Equation Model. Best fit indices presented bolded and underlined. ESEM model presented in italics.

Table 2.

The Generic Conspiracy Believes Scale factor pattern matrixes estimation using bi-geomin, bi-quartimin, a theory-based target and SLiD target.

	Bi-quartimin						Bi-geomin						Theoretical target						SLiD					
	GCI	GM	MG	ET	PW	CI	GCI	GM	MG	ET	PW	CI	GCI	GM	MG	ET	PW	CI	GCI	GM	MG	ET	PW	CI
I1	**.78**	**.34**	-.03	-.07	-.03	.11	**.76**	**.39**	.00	-.04	-.01	.09	**.74**	**.43**	.03	-.03	.01	.09	**.79**	**.34**	-.02	-.05	-.07	.07
I6	**.80**	**.22**	.03	-.05	.06	-.01	**.79**	**.26**	.05	-.03	.07	.00	**.78**	**.28**	.07	-.02	.07	.00	**.81**	**.20**	.04	-.03	-.01	-.02
I11	**.79**	.18	-.04	-.11	-.08	-.06	**.78**	**.25**	-.03	-.10	-.08	.15	**.76**	**.30**	.00	-.08	-.08	.15	**.78**	**.26**	.01	-.06	.05	.19
I2	**.73**	.09	**.40**	-.03	.00	-.05	**.72**	.10	**.41**	-.02	.01	-.01	**.71**	.11	**.43**	-.01	.02	.00	**.72**	.08	**.42**	.00	-.05	-.01
I7	**.74**	-.04	**.45**	.03	.03	.15	**.75**	-.04	**.45**	.03	.03	-.01	**.74**	-.03	**.45**	.03	.02	.01	**.73**	-.04	**.47**	.06	-.01	.01
I12	**.81**	-.04	**.42**	-.02	-.02	.02	**.81**	-.02	**.42**	-.02	-.03	-.05	**.80**	.00	**.43**	-.02	-.04	-.03	**.78**	.02	**.47**	.03	.06	.01
I3	**.65**	-.01	.03	**.61**	.03	.00	**.65**	-.03	.03	**.61**	.04	-.06	**.65**	-.03	.03	**.60**	.03	-.05	**.63**	-.03	.05	**.63**	.00	-.07
I8	**.63**	.04	-.02	**.67**	.01	.04	**.63**	.02	-.02	**.67**	.02	.04	**.63**	.03	-.01	**.67**	.02	.05	**.62**	.01	-.02	**.68**	-.07	.01
I13	**.70**	-.11	-.02	**.52**	-.03	-.11	**.70**	-.10	-.03	**.51**	-.03	-.03	**.70**	-.08	-.02	**.51**	-.06	-.02	**.65**	-.04	.04	**.57**	.12	.02
I4	**.80**	-.01	.02	.05	**.46**	**.29**	**.81**	.00	.01	.04	**.45**	-.08	**.83**	-.02	-.01	.02	**.40**	-.07	**.87**	**-.21**	-.05	.01	-.01	**-.23**
I9	**.76**	-.18	.05	.12	.00	.17	**.77**	-.14	.03	.11	-.02	-.13	**.77**	-.11	.04	.11	-.08	-.12	**.71**	-.04	.15	.19	**.28**	-.03
I14	**.81**	-.12	-.08	.00	.02	-.04	**.82**	-.07	-.09	-.01	.01	.01	**.82**	-.02	-.08	-.01	-.05	.02	**.78**	-.01	.00	.06	**.25**	.09
I5	**.71**	-.06	-.03	.05	.16	-.03	**.72**	-.04	-.04	.03	.15	.13	**.72**	-.01	-.05	.03	.11	.14	**.73**	-.10	-.04	.05	.06	.10
I10	**.60**	.00	.02	.08	-.01	.04	**.61**	.02	.01	.07	-.01	**.27**	**.60**	.06	.02	.08	-.03	**.28**	**.61**	.00	.02	.09	-.02	**.27**
I15	**.69**	.05	-.03	-.07	-.04	**.48**	**.70**	.09	-.04	-.08	-.03	**.45**	**.69**	.15	-.02	-.08	-.06	**.46**	**.71**	.05	-.04	-.06	-.05	**.45**

Note. GCI: General Conspiracist Ideation. GM: Government Malfeasance; MG: Malevolent Global Conspiracies; ET: Extraterrestrial Cover-up; PW: Personal well-being; CI: Control of Information. SLiD: Empirical Iterative Target Rotation based in a Schmid-Leiman Solution. Factor Loadings with an absolute value over .20 are presented bolded and shadowed in grey.

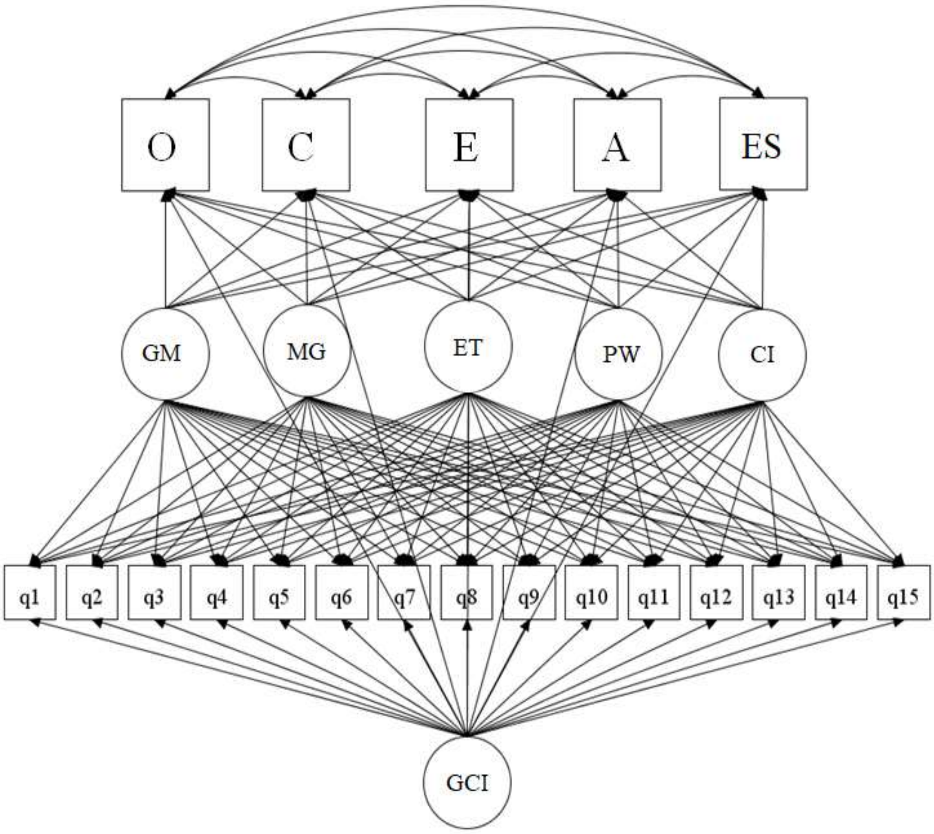

Figure 3. The Bi-factor ESEM model for between the Generic Conspiracist Beliefs Scale scores and personality traits. Items are represented from q1 to 15. GCI: General Conspiracist Ideation. GM: Government Malfeasance; MG: Malevolent Global Conspiracies; ET: Extraterrestrial Cover-up; PW: Personal well-being; CI: Control of Information; O = Openness; C = Conscientiousness; E = Extraversion; A = Agreeableness; ES = Emotional Stability.

Table 3.

Standardized regression coefficients between big five personality traits and CGBS scale.

	GCI	GM	MG	ET	PW	CI	R^2
Openness	**.106**	.006	.037	.002	.072	-.067	.022
Conscientiousness	.020	**-.107**	-.018	.063	**-.138**	.052	.038
Extraversion	.036	-.059	.033	**.078**	-.033	.028	.013
Agreeableness	**-.058**	-.051	**.079**	-.009	-.036	.012	.014
Emotional Stability	**-.076**	.011	**-.078**	**.107**	-.064	.082	.034

Note. GCI: GCI: General Conspiracist Ideation. GM: Government Malfeasance; MG: Malevolent Global Conspiracies; ET: Extraterrestrial Cover-up; PW: Personal well-being; CI: Control of Information. SLiD: Empirical Iterative Target Rotation based in a Schmid-Leiman Solution. Significant regression parameters (at .01 level) presented bolded and shadowed in grey.

Appendix 1

Table A1.

Standardized regression coefficients between big five personality traits and GCBS scale for different bi-factor rotation methods.

GCBS	Personaliy Trait	Bi-quartmin	Bi-geomin	Theory Target	SLiD-based
GCI	Openness	**.095**	**.096**	**.094**	**.106**
	Conscientiousness	.016	.020	.030	.020
	Extraversion	.042	.047	.049	.036
	Agreeableness	**-.056**	-.054	-.049	**-.058**
	Emotional Stability	-.054	**-.056**	**-.058**	**-.076**
GM	Openness	.044	.033	.036	.006
	Conscientiousness	-.093	**-.094**	**-.115**	**-.107**
	Extraversion	-.071	-.076	-.077	-.059
	Agreeableness	-.047	-.057	-.065	-.051
	Emotional Stability	-.047	-.039	-.038	.011
MG	Openness	.024	.024	.048	.037
	Conscientiousness	.060	.057	-.028	-.018
	Extraversion	.007	.064	.020	.033
	Agreeableness	-.008	-.012	**.075**	**.079**
	Emotional Stability	.070	.071	**-.096**	**-.078**
ET	Openness	.048	.048	.024	.002
	Conscientiousness	-.013	-.020	.047	.063
	Extraversion	.028	.021	.062	**.078**
	Agreeableness	**.082**	**.078**	-.017	-.009
	Emotional Stability	**-.100**	**-.099**	**.076**	**.107**
PW	Openness	.008	.014	.021	.072
	Conscientiousness	**.094**	.084	.065	**-.138**
	Extraversion	.001	-.005	-.014	-.033
	Agreeableness	.013	.011	.006	-.036
	Emotional Stability	-.086	-.055	-.059	-.064
CI	Openness	**.095**	**.095**	**.094**	-.067
	Conscientiousness	**-.131**	**-.134**	**-.129**	.052
	Extraversion	-.018	-.018	-.012	.028
	Agreeableness	-.039	-.036	-.034	.012
	Emotional Stability	**-.109**	**-.108**	**-.104**	.082

Note. GCI: Generic Conspiracist Ideation. GM: Government Malfeasance; MG: Malevolent Global Conspiracies; ET: Extraterrestrial Cover-up; PW: Personal well-being; CI: Control of Information. SLiD: Empirical Iterative Target Rotation based in a Schmid-Leiman Solution. Significant regression parameters (at .01 level) presented bolded.

Chapter 7

General discussion

7.1 Main Results

The bi-factor model has emerged as one of the most influential statistical tools in psychological research (Markon, 2019; Reise et al., 2018). Unfortunately, the bi-factor model is not only poorly understood (Bonifay et al., 2017; Markon, 2019), but also commonly applied within a detrimental confirmatory factor analysis approach (Reise et al., 2018, 2011). This doctoral dissertation has been dedicated to the unique challenges associated with approximating the bi-factor model from an exploratory perspective. Particularly, this thesis considered the target rotation as the principal tool for conducting bi-factor exploratory factor rotation. Thus, this doctoral dissertation was focused on developing new target-based rotation algorithms to improve the estimation of exploratory bi-factor models. These algorithms successfully recovered complex bi-factor exploratory factor models under several conditions when explored using Monte Carlo methods and empirical datasets. Moreover, by presenting these algorithms within a unified framework (the SLi family of rotations), it is expected that future contributions and collaborations in this area are set to be continued[1].

In detail, Chapters 2 and 3 were devoted to the introduction of the SLi and SLiD algorithms as state-of-the-art tools for conducting bi-factor exploratory factor analysis. Chapter 4 presented an application of these methods to the evaluation of the factor structure of the Last Twelve matrices of the Standard Progressive Matrices test (i.e., SPM-LS). Chapter 5 included a detailed evaluation of how choosing between different algorithms could ultimately impact the quality of omega hierarchical estimation. Lastly, in Chapter 6, an application was developed so easily translate the SLiD estimated target from R to *Mplus*. Its use and

[1] A first example is an extensive correspondence established with Dr Stephen Robertson during his studies on factor collapse in bi-factor models. In this note, I would like to express my deepest gratitude for sharing his knowledge and enthusiasm during the past years.

utility were exemplified by presenting a novel bi-factor ESEM exploration of the relationship between the Generic Conspiracionist Belief Scale scores and personality traits.

To conclude, it should be acknowledged that this doctoral dissertation intended to shed light on some of the most controversial topics surrounding exploratory bi-factor models. In addition to the comparison of algorithms and the novelty of some of the strategies presented in this thesis, previous chapters also included noteworthy findings that should not be forgotten: novel bi-factor taxonomies and conceptualizations, new approximations towards identifying substantive effects on bi-factor model recovery and the presentation of relevant advances in how the target rotation should be conducted. For the sake of clarity, the main results of each Chapter are hereafter presented.

7.1.1 Main Results from Chapter 2

Chapter 2 illustrated how the iterative application of the partially specified target rotation, as included in the SLi algorithm, represented substantial benefits when conducting bi-factor exploratory factor analysis. This conclusion was supported by evidence obtained from Monte Carlo methods and the examination of an empirical example. The simulation presented a set of comprehensive conditions, including different sample sizes, different number of group factors, different number of indicators per factor and two types of disturbances of the simplest bi-factor model (namely, cross-loadings and true-near zero loadings on the group factors). SLi was shown to outperform other algorithms under most conditions, including structures where group factor structure presented a cross-loadings, or, more importantly, pure indicators. Accordingly, researchers were recommended to use the SLi algorithm.

This study also explored the performance of other available algorithms: bi-quartimin, bi-geomin, SL and non-iterative SL target rotation: Bi-quartimin failed when cross-loadings were present. SL did never provided accurate recovery of bi-factor models presenting pure indicators of the general factor. Bi-geomin showed the best performance when a combination of cross-loadings and near-zero loadings were present in the structure. Lastly, no method performed correctly under small sample sizes (i.e., N = 200) and when group factors had small average factor loadings (i.e., between .30 and .50).

It should be highlighted that this research was one of the first studies comparing the performance of different bi-factor rotation methods using Monte Carlo simulations. As such, it has been already cited 13 times in the literature as of April 2020, and the conditions studied there inspired posterior studies in the field (Giordano and Waller, 2019; Lorenzo-seva and Ferrando, 2018; Robertson, 2019). Additionally, this article presented a terminology inspired in Mcdonald (1999) for distinguishing bi-factor structures including pure indicators:

the independent cluster pure structures (i.e., ICP) and the independent cluster basis pure structures (i.e., ICBP).

7.1.2 Main Results from Chapter 3

Chapter 3 introduced the logic underlying the Empirical Iterative Target Rotation based on a Schmid-Leiman solution algorithm (i.e., SLiD). This Chapter detailed how this algorithm adapted different strategies from the target rotation literature to develop a novel empirical method for establishing factor-specific cut-offs. This method was based on finding jumps in the sorted absolute distribution of loadings' differences for each factor. Ultimately, and contrarily to other cut-offs methods available in the literature, it was explained why the SLiD algorithm did not explicitly aim to maximize the hyperplane count for each factor but to avoid incorrectly fixing substantive cross-loadings to zero in the final partially specified target matrix.

The SLiD algorithm was compared to the original SLi algorithm employing a set of different single, fixed cut-off point (from .05 to .20 in .05 jumps). To provide a deeper depiction of these algorithms functioning, a Monte Carlo simulation was designed. The simulated structures were intended to reproduce realistic conditions where: (a) an extensive number of cross-loadings simulated from normal distributions of different variance were included for each factor in the structure; and (b) a mixture of strong, middle and weak group factors (i.e., group factor differing in their average substantive factor loadings) were present within the same bi-factor structure. Additionally, independent variables considered in Chapter 2 were included in the simulation, alongside with the reanalysis of the empirical example considered in that study.

Results showed that SLiD outperformed SLi with a single, fixed cut-off points. Moreover, SLiD allowed the correct recovery of factors showing a high number of substantial cross-loadings under hindering conditions. Accordingly, researchers were suggested to avoid deciding for a fixed cut-off point and to opt for applying the SLiD algorithm, which was shown to outperform another alternative under most circumstances, albeit examining the quality of the estimated solution. For a few conditions, SLiD could result in factor collapse and unstable performance. This result was concordant with the simulation scheme applied, where many minor cross-loadings were present in the structures with weak group factors. Lastly, it was observed that group factors with low average factor loadings were not correctly recovered, particularly if larger cross-loadings were present. Thus, no method was able to recover the original structure under such detrimental conditions.

7.1.3 Main Results from Chapter 4

Chapter 4 illustrated the usefulness of the methods investigated in Chapter 2 and 3 in the context of an intelligence test, the Last Twelve matrices of the Standard Progressive Matrices (i.e., SPM-LS). In this case, the assumption of essential unidimensionality, crucial in the statistical methods applied in the original validation of the test, was thoroughly tested. Even though traditional and modern dimensionality assessment methods (i.e., parallel analysis and exploratory graph analysis) suggested the presence of relevant sources of variability, the application of modern bi-factor exploratory techniques revealed that SPM-LS scores could be considered as essentially unidimensional. However, results also evidenced that an additional factor emerged for the last six items. Therefore, even though SPM-LS scores could be used as valid measures of individual differences in general intelligence, the correct bi-factor measurement model (including the group factor for the last six items) should be considered in certain contexts due to its better fit to the data.

7.1.4 Main Results from Chapter 5

Chapter 5 provided a detailed comparison of different bi-factor rotation algorithms when estimating the omega hierarchical statistic. In this study, two new methods were examined: the Direct Schmid-Leiman and the Direct Bi-factor methods (Giordano and Waller, 2019; Waller, 2017). All methods were tested over three different Monte Carlo simulations investigating the omega hierarchical recovery under full-rank bi-factor models, second-order models (transformed employing the SL orthogonalization) and a final simulation studying how algorithms would behave when no general factor was present in the structure. Additionally, several accuracy statistics were considered, namely the mean absolute error (i.e., MAE), mean bias error (i.e., MBE) and the root mean square error (i.e., RMSE). Lastly, it should be highlighted that in addition to conditions studied in Chapter 2, the average factor loading of the general factor was manipulated in the first two simulations. Lastly, each algorithm was further evaluated based on the analysis of eight classical datasets presenting full-rank and rank-deficient bi-factor models.

Results suggested that the SLiD algorithm provided the best approximation to omega hierarchical under a majority of conditions and regardless of the type of structure under consideration (full, rank-deficient or no-general factor). However, not even SLiD could provide good recovery of omega hierarchical values under low sample sizes (i.e., N = 150), a low number of factors (i.e., 3) or a small general factor magnitude (i.e., .30). Thus, researchers were warned against trusting omega hierarchical estimates under these conditions. SL, bi-geomin and bi-quartimin did not provide an adequate recovery of omega hierarchical.

Particularly, bi-quartimin and bi-geomin tend to overestimate omega hierarchical, while SL resulted in the opposite behaviour. This result implies that these methods tend to overestimate and underestimate general factor loadings, respectively. While bi-geomin and bi-quartimin general factor overestimation is a well-known effect, SL underestimation was observed for the first time. Such underestimation was attributed to the underestimation of the first-order correlation matrix. Either way, further research must be ensued to explore the ramifications of this finding.

The study of the Direct Schmid-Leiman and the Direct Bi-factor methods revealed that methods based on a completely specified target rotation (based on targets of ones and zeroes) are dependent upon the existence of discrepancies between the average targeted factor loadings of the factors involved in the bi-factor structure. When factors are of similar strength, a fully defined target rotation would provide an unbiased estimation of factor loadings. If factors differ in their average targeted factor loading, rotated factor loadings will be biased towards the common average of the targeted loadings. Furthermore, the small but systematic deviations introduced in the general factor loadings by such effect biased omega hierarchical estimation (regardless of the overall quality of the estimation of these loadings). This study provided evidence that the relationship between the quality of factor loadings estimation and the recovery of secondary statistics is not as straightforward as many researchers could think. Good factor loading recovery could not ensure a correct estimation of these secondary statistics if systematic bias is present. To the author's knowledge, this is a novel result never explicitly stated in the literature with regards to the use of completely and partially specified target rotations.

7.1.5 Main Results from Chapter 6

Chapter 6 was devoted to present the utility of the SLiD algorithm in the context of exploratory structural equation modelling (i.e., ESEM). ESEM are an expansion of traditional SEM to accommodate exploratory measurement models (Marsh et al., 2014, 2009), outperforming the traditional confirmatory alternatives under many realistic conditions (Asparouhov and Muthén, 2009; Guo et al., 2019; Marsh et al., 2019). In this context, bi-factor ESEM models have gained attention in the literature (Morin et al., 2016). Unfortunately, the principal software for conducting ESEM (*Mplus*) only enable users to apply bi-geomin, bi-quartimin or simple target rotation when estimating exploratory bi-factor models. In this chapter, a tutorial on how to perform a SLiD-based bi-factor ESEM was provided, illustrating the features of a novel Shiny application developed to translate the SLiD-based target rotations into a *Mplus* script. This process was exemplified in a novel investigation regarding the structure of the Generic Conspiracionist Believes Scale and its relationships with personality traits. Results showed that SLiD was able to recover unique features of the data when compared to other algorithms. Moreover, it was hoped that by conducting a step-by-step demonstration on how to perform SLiD-based bi-factor ESEM, applied researchers interesting interested in this topic would greatly benefit from using the presented free Shiny app.

7.2 Future Directions and Limitations

7.2.1 The Nature of the Bi-factor Model

First and foremost, it is important to emphasize that, as of today, we have just begun to comprehend the statistical properties of the bi-factor model. Twenty years after the seminal paper by Yung et al. (1999), and a decade after its rediscovery (Reise et al., 2012), there are still plenty of questions surrounding the bi-factor model that remain unresolved (Bonifay et al., 2017; Markon, 2019).

And the first, and most important question yet to be answered is: What is a bi-factor model? Unfortunately, we do not know it yet. In most cases, a bi-factor model is a tag

given to any factor pattern defining a general plus several group factors in which a simple structure is expected to hold. As such, recent articles consider SL solutions to be some sort of a rank-deficient bi-factor model (Waller, 2018; Giordiano & Waller, 2019). Nevertheless, as explained in Chapter 1, this decision could be questionable as not all rank-deficient bi-factor models are consistent with the result of an SL transformation. It might be worthy to distinguish between rank-deficient solutions that are consistent with having being generated from a transformed higher-order model (regardless if they have been obtained employing the SL transformation or any other method) and those which are not. Moreover, the higher-order model and the use of SL transformation as an approach to obtain solutions consistent with the higher-order model should be distinguished in detail. As noted by Gignac (2016): "The higher-order model is a model, for example, as it can be specified and tested statistically for plausibility. The Schmid-Leiman transformation, however, cannot be specified and/or tested for plausibility. It is simply used to calculate indirect effects" (p. 58, footnote 1).

Another relevant debate is whether bi-factor models should be included in the family of hierarchical models (Markon, 2019). Truth is that understanding higher-order and bi-factor models within a common framework could be useful to highlight their similarities and differences (Yung et al., 1999), particularly when researchers are interested in assessing the reasonableness of the constraints implied by each model. In this context, some authors have referred to the higher-order and the bi-factor model as indirect and direct hierarchical models, respectively (Gignac, 2008, 2016). This nomenclature stresses the theoretical implications of each model and the presence of direct or mediated effects of the general factor onto the items. However, some authors have challenged this classification: "[about the bi-factor model] Note that this is not a hierarchical model, because *g* (which is necessarily saturated in all of the variables) does not depend on the variables' loadings on the group factors" (Jennsen & Weng, 1994, p.245). Noteworthy, this debate should not be taken lightly, as the consideration (or not) of the bi-factor model as a hierarchical model could affect its consideration as a plausible measurement model within certain areas of research (i.e., intelligence or psychopathology). Either way, it is beyond doubt that the literature would benefit from agreeing on a unique terminology for referring to bi-factor, second-order, and similar models including one (or more) general factors.

Lastly, some authors have strongly argued that researchers must explicitly specify how the group factor represent subordinate facets of the general factor. To this end, the bi-factor S-1 model was proposed by Eid et al. (2016, 2018). In this model, a set of items (all loading onto the same group factor) act as pure indicators of the general factor. By using the S-1 bi-factor model, the researcher ensures that the general factor represents variance of the facet removed (which acts a reference facet). Moreover, the use of the oblique S-1 factor model could prevent identifiability issues in bi-factor estimation (Eid et al., 2018). This interesting perspective should be explored in greater detail in the future.

7.2.2 Questions in Exploratory Bi-factor Model

As Reise et al. (2018) indicated, there still exist several misconceptions surrounding the bi-factor model: (a) its confirmatory nature; (b) its inability to accommodate disturbances of the simple structure; and (c) the requirement of orthogonal group factors. In this thesis, the first two questions have been addressed and discussed in length. Regarding the latter, and to expand on the matter, it is important to recall that oblique factor solutions should generally be preferred to orthogonal ones, as they enable the recovery of simple solutions in the pattern matrix while providing evidence for the relationship between factors (Browne, 2001). To this end, Jennrich and Bentler (2012) developed bi-geomin for estimating oblique bi-factor models. The oblique bi-factor model might be observed in the presence of method factors (for example, when acquiescence or wording effects are present), or when bi-factor models are applied within the context of multitrait-multimethod or multi-rater analysis (Eid et al., 2018), among other instances (Lorenzo-seva and Ferrando, 2018). Thus, there might be contexts where the oblique bi-factor model could be of merit.

However, researchers should be cautious when applying oblique bi-factor modelling. Firstly, the question of which rotation should be preferred when estimating factor correlations does not have a clear response. Not even in the general, non-bifactor case: while some authors favour geomin (Celimli Alkoy, 2017), recent advances in target rotation remain largely unexplored (Zhang et al., 2018a). Noteworthy, a comparison of these methods in the context of bi-factor models has yet to be presented. Secondly, the accuracy and stability of factor correlations estimation could be particularly inefficient in the bi-factor case, resulting in difficulties with regards to their convergence and the replicability of these models. Thirdly, and more importantly, the introduction of oblique group factors prevents researchers from interpreting bi-factor models as in the orthogonal case. In the traditional bi-factor model, group factors are specified to reflect unique sources of residual variance to the general factor, which is no longer true if group factors correlate. In Reise et al. (2018) words: "By no means does this imply that estimating bi-factor models with correlated group factors are not complicated or that solutions are readily interpretable; they are not" (p.684, footnote 3). This scepticism could underlie why oblique bi-factor models have received little attention so far in the literature.

Lastly, it should be highlighted that all reviewed bi-factor models presents an additional sacrosanct restriction, unavoidable even in exploratory bi-factor modelling: the orthogonality assumption between the general and group factors (Eid et al., 2016; Markon, 2019). Such restriction is required for the model to be identified and estimable (Markon, 2019): "General-specific factor [group factor] correlations are likely to be inadmissible regardless of the scenario" (p.12.10). Accordingly, fully unrestricted bi-factor models (where the general factor present direct effects on both, items and group factors), as in Yung et al. (1999), could be considered more of a mathematical curiosity than a suitable model to be applied in real-world settings. But, as today's limitations could represent tomorrow's opportunities, the extent that these models could be approached should be explored in the future.

7.2.3 The Expanded Bi-factor Model

The bi-factor model has been recently extended to accommodate the presence of different variance sources beyond one general factor and one substantive group factor per item: (a) The two-tier factor models, which presents an additional general factor (Cai, 2010; Cai et al., 2011); and (b) the tri-factor model, where each item presents a substantive loading onto several group factors (Bauer et al., 2013; Jeon et al., 2018). In this context, the amalgam of complex bi-factor models that could be derived in the near future might only grow exponentially. Nevertheless, researchers should proceed with extreme caution when applying these models, understanding their unique benefits, drawbacks and theoretical interpretation. It is also important to bear in mind that the higher the complexity, the higher the number of restrictions needed for the model to be identified (and the higher the chances that some of these restrictions do not hold). Due to their highly structured nature, these models are based on confirmatory approaches, where the researcher is often required to impose key constrains for ensuring their convergence (i.e., in a two-tier model, general factors can correlate between them, but not with group factor indicators; Cai, 2010). Thus, these alternative models might benefit from being considered from an exploratory (unrestricted) perspective. Lastly, it is still to be seen how to integrate these models within the aforementioned taxonomies applied in bi-factor modelling.

7.2.4 The Plausibility of the Bi-factor Model

One of the main appeals of the bi-factor model is that it allows researchers to evaluate the presence of a general factor which presents direct effects on the items. This doctoral dissertation (Chapter 5) focused on studying omega hierarchical as a tool for such a task. This index currently plays a crucial role when researchers understand the quality of the total scores derived from a bi-factor model. As such, it is routinely reported in most bi-factor publications. For example, omega hierarchical has been a focal point of the debate regarding whether a general factor of personality is reasonable or not (Arias et al., 2018; Revelle and Wilt, 2013). Omega hierarchical, as well as other secondary statistics (including the explained common variance or the H-index), are only expected to grow in importance as bi-factor models are applied in more research contexts (Chen and Zhang, 2018; Reise et al., 2018; Rodriguez et al., 2016).

However, a note of caution should be sounded here regarding the use of general factor modelling in contexts where their theoretical status is, to say at least, dubious (Borsboom et al., 2003; van Bork et al., 2017; van der Maas et al., 2006). Researchers must understand that under common circumstances found in psychological testing, where the *positive manifold* hypothesis[2] is expected to hold, a general factor could be observed regardless of the true generative mechanism of the observed correlations (Borsboom and Wijsen, 2017). Indeed, extracting a general factor from a set of positive correlation is only a necessary, but not sufficient proof of its "existence", as established by Perron-Frobenius theorem: under a covariance matrix of positive entries, there exists a positive eigenvalue (of largest value) corresponding to a single positive eigenvector whose entries are all positive[3]. In this sense, extracting a general factor provides little to none information of the true cause of the positive variance-covariance matrix. Unfortunately, while the Perron-Frobenius theorem has been

[2]The positive manifold hypothesis, which refers to psychological tests being invariably positively correlated, was discussed since the early days (Spearman, 1904; van der Maas et al., 2006).

[3]For its adaptation to the single and hierarchical simplest structure factor analysis, see Theorems 1 and 2 in Krijnen (2004).

used to justify the unfalsifiability of the common factor model (van Bork et al., 2017), Borg (2018) showed that only holds if the dimensionality of the common space is, at most, two. Under larger common spaces, the positive manifold is no longer a sufficient, but a necessary condition for a general factor (a factor with all positive and substantial loadings) to occur.

Either way, these considerations should act as warning signals that any substantial decision based on general factor modelling (and bi-factor models, particularly) should be based on a strong statistical *and theoretical* foundations, and not to solely rely on the results of factor analysis (or other correlation-based methods). As Thurstone reflected: "The exploratory nature of factor analysis is often not understood. Factor analysis has its principal usefulness at the border of science" (Thurstone, 1947, p.56)[4]. Indeed, pretending that factor analysis to go beyond these frontiers could be a rather unfruitful and rather overoptimistic task. Unfortunately, Thurstone's cautionary tale with regards to exploratory factor analysis was lost in decades of mechanical application of this statistical technique [5].

The evaluation of group factors is well behind that of the general factor, particularly in exploratory bi-factor modelling. Therefore, it is imperative to develop appropriate alternatives of indices such as omega hierarchical for group factors in exploratory bi-factor research, particularly given the challenges observed in the estimation, interpretation and evaluation of group factor loadings and scores (Bonifay et al., 2017; Markon, 2019; Reise et al., 2013). As of today, there exist rightful concerns about whether group factors have a significant role in bi-factor modelling as well as the nature of the variance they account for (Bonifay et al., 2017; Sellbom and Tellegen, 2019). By definition, in (orthogonal) bi-factor modelling groups factor reflect variance sources residual and distinct from the general factor. However, this assumption is questioned on many applications of the bi-factor model. As a result, the evaluation of the relationship between general and group factors is a challenging task with no simple solution so far (Eid et al., 2018; Markon, 2019). Nevertheless, without improving this aspect, the discriminant and convergent validity of the bi-factor model could be compromised.

Obtaining reliable estimations of the group factors parameters is also quite problematic. Group factors are often defined by just a few indicators presenting an amalgam of factor loadings of different magnitude. This situation results in group factors scores being unstable and unreliable. In their systematic review, Rodriguez et al. (2015) found that while the general

[4]Thurstone demonstrated many of its advances on factor analysis using physical quantities, such as cylinders shapes or the famous box measurements problems. Unfortunately, such a tradition is nowadays lost. Such drift has driven factor analysis from its mathematical roots to more direct applications, sometimes without the necessary comprehension of its statistical properties under standardized conditions.

[5]Noteworthy, other prominent authors were far less careful when explaining the limitations of exploratory factor analysis: "[...] the existence of distinct nebulae of variables in the correlational configurations from natural data is no more an accident than the existence of nebulae or the Milky Way in the sky. Laws of nature brought such structures about in both cases." (Cattell, 1978, p.105)

factor was represented with an average of almost 20 items, group factors were estimated using an average of only 7 items. Additionally, while the average omega hierarchical was .80, while average omega hierarchical subscale was as low as .27. Similar indices have also been observed in of psychopathology scales (Constantinou and Fonagy, 2019). Thus, the question of the usefulness of the group factors beyond general factor scores has been debated in the literature (Bonifay et al., 2017; Rodriguez et al., 2016).

Additionally, the estimation of group factors could be diminished by a more subtle consequence of the presence of general factor: the tendency for the general factor (in a bi-factor model) to accommodate meaningless patterns of responses (Reise et al., 2016; Watt et al., 2019). This situation reveals that a general factor will absorb both, substantial and unsubstantial common variance, regardless of the researcher's intention or model specification. For example, under the presence of a strong method's factor underlying a majority of items of a given scale (e.g., acquiescence), it is unclear what the general factor would stand for. While this situation opens bi-factor models to be of utility when controlling for such method factors, it also raises questions regarding what does the general factor represents when these method artefacts cannot be explicitly controlled. Similarly, the bi-factor model has been argued to present a better data fit than alternative models in many instances where it should not be the case (Bonifay and Cai, 2017; Canivez, 2016; Cucina and Byle, 2017; Murray and Johnson, 2013; Rodriguez et al., 2016). The reasons underlying that propensity to present better fit has been connected with the rank restrictions present in other models such as the higher-order model (Gignac, 2016; Molenaar, 2016; Yang et al., 2017) and the (questionable) ability for the bi-factor model to accommodate implausible data patterns. The extent these issues continue to occur in exploratory bi-factor models constitutes an intriguing future line of work.

7.2.5 Challenges in Factor Rotation

The results presented in this doctoral dissertation have supported the application of the iterative partially specified target rotation as a reliable method for approximating complex structures such as the bi-factor model. One of the principal results of this dissertation was the development of the SLiD algorithm: a novel method for finding suitable factor-specific, empirical cut-off points for defining the target matrix, and that resulted in improved factor recovery when compared with traditional schemes of using single cut-off points. Unfortunately, while there exist several different approaches towards finding empirically-defined partially specified target rotation, namely Promaj (Trendafilov, 1994) or Promin (Lorenzo-seva, 1999; Lorenzo-Seva and Ferrando, 2019a), to mention a few, these have been scarcely compared in the factor rotation literature, not to say in the context of bi-factor exploratory factor analysis. To say that during the latest stages of this dissertation, a new algorithm called Pure Exploratory Bifactor Modelling (i.e., PEBI) was proposed in the literature (Lorenzo-seva and Ferrando, 2018). In short, this algorithm employs a Promin-based cut-off point definition to distinguish between near-zero and relevant value in the target matrix. PEBI is a promising approach that expands the capabilities of partially specified target rotation methods to the oblique group factor and the single group factor case. Readers are encouraged to test the PEBI algorithm (as well as the other unique capabilities) of the fantastic and free FACTOR program for conducting factor analysis (Ferrando and Lorenzo-Seva, 2017b), which I cannot personally recommend enough.

Target rotation is an exciting area of research within psychometrics. Its mathematical simplicity, coupled with its ability to recover factor structures under diminished conditions (Browne, 2001; Fleming, 2012), ensures that it will continue playing a substantive role in the estimation of complex factor structures inside and outside the bi-factor case (Guo et al., 2019; Marsh et al., 2019). As of today, target rotation could be expected to constitute the cornerstone of future exploratory approaches towards more complex factor structures (such as two-tier or tri-factor models). Several expansions of traditional target methods have been discussed in the literature with limited success, and are waiting for interested researchers to be applied within or outside the bi-factor context: from understanding the consequences of specifying targets into other parts of the factor model (Zhang et al., 2018b), to the application of bootstrapping techniques to understand the stability of the rotation estimation (Paunonen, 1997) or to develop novel schemes for target rotation refinement (Lorenzo-seva and Ferrando, 2018). But before advancing on new any new area of target rotation, more attention should be drawn to the comparison of the most widely used target methods: the completely specified and the partially specified target rotation. As the decision of choosing one way or the other could shape future research in the area, the benefits and drawbacks must be carefully examined in both, future simulation and empirical studies.

Lastly, it should be kept in mind that the target rotation is nothing but a small portion of a wider array of mathematical models falling into the Generalized Procrustean Analysis umbrella (Crosilla et al., 2019). Noteworthy, while some of the most relevant advances on target rotation have been produced in psychometrics, key contributions have been produced in areas as shape analysis, computer vision, etc. (Gower and Dijksterhuis, 2004). As the connection between these research areas has been rather scarce, some of the problems observed in factor rotation might have been addressed in their literature. As an example, the large number of alternatives for conducting partially specified target rotation available in Crosilla et al. (2019, p.20), and their benefit remain largely unknown by our research community.

The interest in exploratory solutions is fostering advances in factor rotation, with many alternative approaches being published in major journal and publication outlets. Even though some authors might believe that this old topic is largely resolved in psychometrics (Mulaik, 1986), there are many areas of improvement: (a) to gain a deeper understanding of the available minimization methods, such as gradient projection algorithm functioning. For example, by confirming its statistical behaviour when compared with other optimization algorithms (Weide and Beauducel, 2019). In this sense, little is known regarding the sensibility of different parameters fixed in this algorithm (the learning rate, etc.). As an example, the use of random oblique transformation matrices (instead of orthogonal ones, as suggested by Mulaik, 2010, p.363) has been confirmed in (unreported) analysis performed by the author of this dissertation to have a substantial impact on the minimization process; (b) to explore different minimization algorithms so to improve rotation estimation in complex criteria (as bi-geomin) or to compare the use of closed-form solutions in target rotation vs the solution found by the gradient projection algorithm; (c) the use of rotations not focused on finding a simple structure concept or hyperplane count. For example, Jennrich (2004b, 2006) proposal of the Component-Loss Criteria, which aims to minimize the absolute value of a given loading, constitute an interesting approach that should be explored in future bi-factor research. On the contrary, understanding rotations such as Varimin (Ertel, 2013), which aim to maximize complexity, would allow researchers to understand the limits of the rotation procedures; and (d) to ensure the identification of the rotated solutions by employing strategies based on the use of the Fisher information matrix (Asparouhov and Muthén, 2009) or similar.

Lastly, factor rotation might be seen its last days as we currently know it. Similar to the transition from indirect to direct rotation methods, a plethora of new methods for conducting these analyses are gaining traction in mainstream psychometrics. Firstly, Bayesian approaches, which rely on the use of small-variance priors in a semi-confirmatory fashion are being strongly considered as an alternative to simple structure-based rotation methods (Asparouhov and Muthén, 2009). However, it is yet unclear which method should be preferable under which conditions, particularly given the sensibility of these Bayesian methods to misspecification problems (Guo et al., 2019; Marsh et al., 2009; Xiao et al., 2019). Be that as it may, at the end of the day both approaches could benefit from each other and to be complementary (Moore et al., 2015). Secondly, and based on the widespread application of machine learning methods, the first estimation on parameter regularization[6] have started to appear in the field (Scharf and Nestler, 2019b; Yamamoto et al., 2017). This approach, which again represents another perspective on removing factor indeterminacy, has several

[6]For interested readers, Jennrich's Component Loss Rotation is inherently related with regularization methods proposed in factor analysis (Scharf and Nestler, 2019a).

benefits, such as having penalization as a parameter to be estimated, and to protect researchers from overfitting (Goretzko et al., 2019; Scharf and Nestler, 2019b). Regularization could represent a perspective worthy of future exploration. Lastly, there has been some interesting reformulations of the factor model itself, aimed to supersede some of the limitations present in the original proposal by overcoming some of the reviewed indeterminacies (Adachi and Trendafilov, 2018, 2019; Sočan, 2003; Stegeman, 2016).

7.3 Conclusion

This doctoral dissertation should be concluded with the same spirit that it was started, and that is by providing a critical view of one of the most influential advances in psychometrics history: the bi-factor model. As of today, it should surprise nobody that the bi-factor model constitutes an indispensable model within the factor analyst toolbox. As such, the bi-factor model is expected to continue growing in relevance, playing an undeniable role in shaping the future of major areas of psychological science. Thus, its strengths and limitations must be ensured to be correctly understood if we aim to have a pertinent use of this statistical tool. At heart, this doctoral dissertation was devoted to delving in our knowledge of this model. And by doing so, it is hoped not only to have realized limited but meaningful contributions in the field but to have inspired others to continue inquiring on the nature and contributions of the bi-factor model.

References

Abad, F. J., Garcia-Garzon, E., Garrido, L. E., and Barrada, J. R. (2017). Iteration of Partially Specified Target Matrices: Application to the Bi-Factor Case. *Multivariate Behavioral Research*, 52(4):416–429.

Adachi, K. and Trendafilov, N. T. (2018). Some Mathematical Properties of the Matrix Decomposition Solution in Factor Analysis. *Psychometrika*, 83(2):407–424.

Adachi, K. and Trendafilov, N. T. (2019). Some inequalities contrasting principal component and factor analyses solutions. *Japanese Journal of Statistics and Data Science*, 2(1):31–47.

Aguirre-Urreta, M. I., Rönkkö, M., and McIntosh, C. N. (2018). A Cautionary Note on the Finite Sample Behavior of Maximal Reliability. *Psychological Methods*, 24(2):236–252.

Anderson, T. W. and Rubin, H. (1956). Statistical Inference in Factor Analysis. In *Proceedings of the Third Berkeley Simposium on Mathematical Statistics and Probability, Volume 5: Contributions to Econometrics, Industrial Research, and Psychometry*, pages 111–150, Berkeley, California. University of California Press.

Arias, V. B., Jenaro, C., and Ponce, F. P. (2018). Testing the Generality of the General Factor of Personality: An Exploratory Bifactor Approach. *Personality and Individual Differences*, 129:17–23.

Asparouhov, T. and Muthén, B. (2009). Exploratory Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(3):397–438.

Asparouhov, T. and Muthén, B. (2019). Nesting and Equivalence Testing for Structural Equation Models. *Structural Equation Modeling*, 26(2):302–309.

Bandalos, D. H. (2018). *Measurement Theory and Applications for the Social Sciences*. Guilford Press, Boca Raton, FL.

Bauer, D. J., Howard, A. D., Baldasaro, R. E., Curran, P. J., Hussong, A. M., Chassing, L., and Zucker, R. A. (2013). A Tri-Factor Model for Integrating Rating Across Multiple Informants. *Psychological Methods*, 18(4):475–493.

Beauducel, A. and Hilger, N. (2019). Score Predictor Factor Analysis: Reproducing Observed Covariances by Means of Factor Score Predictors. *Frontiers in Psychology*, 10(1895).

Bekker, P. A. (1997). Generic Global Indentification in Factor Analysis. *Linear Algebra and its Applications*, 264:255–263.

Bonifay, W. and Cai, L. (2017). On the Complexity of Item Response Theory Models. *Multivariate Behavioral Research*, 52(4):465–484.

Bonifay, W., Lane, S. P., and Reise, S. P. (2017). Three Concerns With Applying a Bifactor Model as a Structure of Psychopathology. *Clinical Psychological Science*, 5(1):184–186.

Borg, I. (2018). A Note on the Positive Manifold Hypothesis. *Personality and Individual Differences*, 134(1):13–15.

Borsboom, D., Mellenbergh, G. J., and Van Heerden, J. (2003). The Theoretical Status of Latent Variables. *Psychological Review*, 110(2):203–219.

Borsboom, D. and Wijsen, L. D. (2017). Psychology's atomic bomb. *Assessment in Education: Principles, Policy and Practice*, 24(3):440–446.

Browne, M. W. (1972). Orthogonal Rotation to a Partially Specified Target. *British Journal of Mathematical and Statistical Psychology*, 25:115–120.

Browne, M. W. (2001). An Overview of Analytic Rotation in Exploratory Factor Analysis. *Multivariate Behavioral Research*, 36(1):111–150.

Cai, L. (2010). A Two-Tier Full-Information Item Factor Analysis Model with Applications. *Psychometrika*, 75(4):581–612.

Cai, L., Yang, J. S., and Hansen, M. (2011). Generalized Full-Information Item Bifactor Analysis. *Psychological Methods*, 16(3):221–248.

Canivez, G. L. (2016). Bifactor Modeling in Test Construction Validation of Multifactored Tests: Implications for Understanding Multidimensional Constructs and Test Interpretation. In Schweizer, K. and DiStefano, editors, *Principles and methods of test construction: Standards and recent advancements*, pages 247–271. Hogrefe, Gottingen, Germany.

Carroll, J. B. (1953). An Analytical Solution for Approximating Simple Structure in Factor Analysis. *Psychometrika*, 18(1):22–38.

Carroll, J. B. (1957). Biquartimin Criterion for Rotation to Oblique Simple Structure in Factor Analysis. *Science*, 126(3283):1114–1115.

Casper, A. and Gower, J. (2010). A General Approach to Handling Misssing Values in Procrustes Analysis. *Advances in Data Analysis and Classification*, 4(4):223–237.

Caspi, A. and Moffitt, T. E. (2018). All for one and one for all: Mental disorders in one dimension. *American Journal of Psychiatry*, 175(9):831–844.

Catell, R. B. (1966). The Data Box: Its Ordering of Total Resources in Terms of Possible Relational Systems. In Catell, R. B., editor, *Handbook of Multivariate Experimental Psychology*. Rand-McNally, Chicago, first edition edition.

Cattell, R. B. (1978). *The Scientific Use of Factor Analysis in Behavioral and Life Sciences*. Plenum, New York.

Cattell, R. B. and Muerle, J. H. (1960). The "Maxplane" Program for Factor Rotation to Oblique Simple Structure. *Educational and Psychological Measurement*, 20(3):569–590.

Celimli Alkoy, S. (2017). *A Comparison of Geomin versus Target Rotation Criteria in Exploratory Factor Analysis with Correlated Factors and Large and Complex Pattern Matrices*. PhD thesis, University of Miami.

Chen, F. F., West, S. G., and Sousa, K. (2006). A Comparison of Bifactor and Second-Order Models of Quality of Life Fang. *Multivariate Behavioral Research*, 41(2):147–163.

Chen, F. F. and Zhang, Z. (2018). Bifactor Models in Psychometric Test Development. In Irwing, P., Booth, T., and Hughes, D. J., editors, *The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale and Test Development*. John Wiley & Sons, Ltd.

Cliff, N. (1966). Orthogonal Rotation to Congruence. *Psychometrika*, 31(1):33–42.

Constantinou, M. P. and Fonagy, P. (2019). Evaluating Bifactor Models of Psychopathology Using Model-Based Reliability Indices.

Crosilla, F., Beinat, A., Fusiello, A., Maset, E., and Visintini, D. (2019). *Advances Procrustes Analysis Models in Photogrammetric Computer Vision*. Springer Nature Switzerland AG, Cham, Switzerland.

Cucina, J. and Byle, K. (2017). The Bifactor Model Fits Better Than the Higher-Order Model in More Than 90% of Comparisons for Mental Abilities Test Batteries. *Journal of Intelligence*, 5(3):27.

Cudeck, R. and O'Dell, L. L. (1994). Applications of Standard Error Estimates in Unrestricted Factor Analysis: Significance Tests for Factor Loadings and Correlations. *Psychological Bulletin*, 115(3):475–487.

Cureton, E. E. and Mulaik, S. A. (1971). On Simple Structure and the Solution to Thurstone's Invariant Box Problem. *Multivariate Behavioral Research*, 6(4):375–387.

Derflinger, G. and Kaiser, H. F. (1989). Method Horst-Hilsch (HH) for the Transformation Problem in Exploratory Factor Analysis. *Psychological Reports*, 65(1):211–220.

Dun, J. E. (1973). A Note on a Sufficiency Condition for Uniqueness of a Restricted Factor Matrix. *Psychometrika*, 38(1):141–143.

Eber, H. W. (1966). Toward Oblique Simple Structure: Maxplane. *Multivariate Behavioral Research*, 1(1):112–125.

Eid, M., Geiser, C., Koch, T., and Heene, M. (2016). Anomalous Results in G-factor Models: Explanations and Alternatives. *Psychological Methods*, 22(3):541–562.

Eid, M., Krumm, S., Koch, T., and Schulze, J. (2018). Bifactor Models for Predicting Criteria by General and Specific Factors: Problems of Nonidentifiability and Alternative Solutions. *Journal of Intelligence*, 6(3):42.

Elden, L. and Trendafilov, N. T. (2017). Towards Determinate Exploratory Factor Analysis.

Epskamp, S., Maris, G. K. J., Waldorp, L. J., and Borsboom, D. (2016). *Network Psychometrics*. PhD thesis, University of Amsterdam.

Ertel, S. (2011). Exploratory factor analysis revealing complex structure. *Personality and Individual Differences*, 50(2):196–200.

Ertel, S. (2013). *Factor Analysis: Healing an Ailing Model*. Universitätsverlag Göttingen, Göttingen.

Ferrando, P. J. and Lorenzo-Seva, U. (2000). Unrestricted versus restricted factor analysis of multidimensional test items: some aspects of the problem and some suggestions. *Psicológica*, 21:301–323.

Ferrando, P. J. and Lorenzo-Seva, U. (2017a). Assessing the Quality and Appropriateness of Factor Solutions and Factor Score Estimates in Exploratory Item Factor Analysis. *Educational and Psychological Measurement*, 78(5):762–780.

Ferrando, P. J. and Lorenzo-Seva, U. (2017b). Program FACTOR at 10: Origins, Development and Future Directions. *Psicothema*, 29(2):236–240.

Ferrando, P. J. and Navarro-González, D. (2018). Assessing the quality and usefulness of factor-analytic applications to personality measures: A study with the statistical anxiety scale. *Personality and Individual Differences*, 123:81–86.

Fleming, J. S. (2003). Computing Measures of Simplictiy of fit for Loadings in Factor-Analytially Derived Scales. *Behavior Research Methods*, 34(4):520–524.

Fleming, J. S. (2012). The Case for Hyperplane Fitting Rotations in Factor Analysis: A Comparative Study of Simple Structure. *Journal of Data Science*, 10:419–439.

Fontanella, L., Fontanella, S., Valentini, P., and Trendafilov, N. T. (2019). Simple Structure Detection Through Bayesian Exploratory Multidimensional IRT Models. *Multivariate Behavioral Research*, 54(1):100–112.

Garnett, J. C. M. (1919). On Certain Independent Factors in Mental Measurements. *Proceedings of the Royal Society of London*, 96(A):91–111.

Gignac, G. E. (2008). Higher-order models versus direct hierarchical models: g as superordinate or breadth factor? *Psychology Science*, 50(1):21–43.

Gignac, G. E. (2016). The Higher-Order Model Imposes a Proportionality Constraint: That is why the Bifactor Model Tends to Fit Better. *Intelligence*, 55:57–68.

Gignac, G. E., Reynolds, M. R., and Kovacs, K. (2017). Digit Span Subscale Scores May Be Insufficiently Reliable for Clinical Interpretation: Distinguishing Between Stratified Coefficient Alpha and Omega Hierarchical. *Assessment*, 26(8).

Giordano, C. and Waller, N. G. (2019). Recovering Bifactor Models: A Comparison of Seven Methods. *Psychological Methods*, pages 1–14.

Goretzko, D., Pham, T. T. H., and Buhner, M. (2019). Exploratory Factor Analysis: Current use, Methodological Developments and Recommendations for Good Practice. *Current Psychology*.

Gorsuch, R. L. (1983). *Factor Analysis*. Lawrence Erlbaum, Hillsdale, N.J., 2nd edition edition.

Gower, J. C. and Dijksterhuis, G. B. (2004). *Procrustes Problem.* Oxford University Press.

Guo, J., Marsh, H. W., Parker, P. D., Dicke, T., Ludtke, O., and Diallo, T. M. O. (2019). A Systematic Evaluation and Comparison Between Exploratory Structural Equation Modelling and Bayesian Structural Equation Modelling. *Structural Equation Modeling: A Multidisciplinary Journal*, 26(4):529–556.

Hancock, G. R. and Mueller, R. O. (2001). Rethinking Construct Reliability Within Latent Variable Systems. In Cudeck, R., Toit, D., and Soerbom, D., editors, *Structural equation modeling: Present and future- A festschrift in honor of Karl Jöreskog*, pages 195–216. Scientific Software International, Lincolnwood.

Harman, H. H. (1967). *Modern Factor Analysis (2d ed.).* University of Chicago Press, Chicago.

Hattori, M., Zhang, G., and Preacher, K. J. (2017). Multiple Local Solutions and Geomin Rotation. *Multivariate Behavioral Research*, 52(6):720–731.

Hayashi, K. and Marcoulides, G. A. (2006). Examining Identification Issues in Factor Analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, 13(4):631–345.

Hendrickson, A. E. and White, P. O. (1964). PROMAX: A Quick Method for Rotation to Oblique Simple Structure. *British Journal of Mathematical and Statistical Psychology*, 17:65–70.

Holzinger, K. J. and Swineford, F. (1937). The Bi-factor Method. *Psychometrika*, 2(1):41–54.

Horst, P. (1941). A Non-Graphical Method for Transforming an Arbitrary Factor Matrix into a Simple Structure Factor Matrix. *Psychometrika*, 6(2):79–99.

Howe, W. G. (1955). Some Contributions to Factor Analysis. *Oak Ridge National Laborayory.*

Hudziak, J., Achenbach, T., Althoff, R., and Pine, D. (2007). A Dimensional Approach to Developmental Psychopathology. *International Journal of Methods in Psychiatric Research*, 16(S1):S16–S23.

Hurley, J. R. and Cattell, R. B. (1962). The Procrustes Program: Producing Direct rotation to Test a Hypothesized Factor Structure. *Behavioral Science*, 7(2):258–262.

Izquierdo, I., Olea, J., and Abad, F. J. (2014). Exploratory Factor Analysis in Validation Studies: Uses and Recommendations. *Psicothema*, 26(3):395–400.

Jennrich, R. I. (2001). A Simple General Procedure for Orthogonal Rotation. *Psychometrika*, 66(2):289–306.

Jennrich, R. I. (2002). A Simple General Method for Oblique Rotation. *Psychometrika*, 67(1):7–19.

Jennrich, R. I. (2004a). Derivative free gradient projection algorithms for rotation. *Psychometrika*, 69(3):475–480.

Jennrich, R. I. (2004b). Rotation to Simple Loadings Using Component Loss Functions: The Orthogonal Case. *Psychometrika*, 69(2):257–273.

Jennrich, R. I. (2006). Rotation to Simple Loadings Using Component Loss Functions: The Oblique Case. *Psychometrika*, 71(1):173–191.

Jennrich, R. I. (2007). Rotation Algorithms: From Beginning to End. In Lee, S., editor, *Handbook of Latent Variable and Related Models*, pages 45–63. Elsevier B.V.

Jennrich, R. I. (2018). Rotation. In Irwing, P., Booth, T., and Hughes, D. J., editors, *The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale and Test Development*. John Wiley & Sons, Ltd.

Jennrich, R. I. and Bentler, P. M. (2011). Exploratory Bi-Factor Analysis. *Psychometrika*, 76(4):537–549.

Jennrich, R. I. and Bentler, P. M. (2012). Exploratory Bi-factor Analysis: The Oblique Case. *Psychometrika*, 77(3):442–454.

Jennrich, R. I. and Sampson, P. F. (1966). Rotation for Simple Loadings. *Psychometrika*, 31(3):313–323.

Jennsen, A. R. and Weng, L. L. (1994). What is a Good g? *Intelligence*, 18(3):231–258.

Jeon, M., Rijmen, F., and Rabe-Hesketh, S. (2018). CFA Models with Factor and Multiple Sets of Secondary Factors. *Psychometrika*, 83(4):785–808.

Jöreskog, K. G. and Lawley, D. N. (1968). New Methods in Maximum Likelihood Factor Analysis. *British Journal of Mathematical and Statistical Psychology*, 21:85–96.

Jöreskog, K. G. (1977). Factor Analysis by Least-Squares and Maximum Likelihood Methods. In Enslein, K., Ralston, A., and Wilf, H., editors, *Mathematical Methods for Digital Computers*, pages 125–153. John Wiley & Sons, Ltd.

Kaiser, H. F. (1958). The Varimax Criterion for Analytic Rotation in Factor Analysis. *Psychometrika*, 23(3):187–200.

Kiers, H. A. L. (1994). Simplimax: Oblique Rotation to an Optimal Target with Simple Structure. *Psychometrika*, 59(4):567–579.

Korth, B. and Tucker, L. R. (1976). Procrustes Matching by Congruence Coefficients. *Psychometrika*, 41(4):531–535.

Krijnen, W. P. (2004). Positive Loadings and Factor Correlations Form Positive Covariance Matrices. *Psychometrika*, 69(4):655–660.

Lawley, D. N. (1958). Estimation in Factor Analysis Under Various Initial Assumptions. *British Journal of Mathematical and Statistical Psychology*, 11:1–12.

Lawley, D. N. and Maxwell, E. (1964). Factor Transformation Methods. *British Journal of Mathematical and Statistical Psychology*, 17:97–103.

Ledermann, W. (1937). On the Rank of the Reduced Correlational Matrixin Multiple-Factor Analysis. *Psychometrika*, 2(2):85– 93.

Loken, E. (2005). Identification Constraints and Inference in Factor Models. *Structural Equation Modeling: A Multidisciplinary Journal*, 12(2):232–244.

Lorenzo-seva, U. (1999). Promin: A Method for Oblique Factor Rotation. *Multivariate Behavioral Research*, 34(3):347–365.

Lorenzo-seva, U. and Ferrando, P. J. (2018). A General Approach for Fitting Pure Exploratory Bifactor Models. *Multivariate Behavioral Research*, 54(1):15–30.

Lorenzo-Seva, U. and Ferrando, P. J. (2019a). Robust Promin: a Method for Diagonally Weighted Factor Rotation. *Liberabit*, 25:99 – 106.

Lorenzo-Seva, U. and Ferrando, P. J. (2019b). Unrestricted factor analysis of multidimensional test items based on an objectively refined target matrix. *Behavior Research Methods*, 52:116–130.

Lorenzo-seva, U. and ten Berge, J. M. F. (2006). Tucker ' s Congruence Coefficient as a Meaningful Index of Factor Similarity. *Methodology*, 2(2):57–64.

Mansolf, M. and Reise, S. P. (2016). Exploratory Bifactor Analysis: The Schmid-Leiman Orthogonalization and Jennrich-Bentler Analytic Rotations. *Multivariate Behavioral Research*, 51(5):698–717.

Markon, K. E. (2019). Bifactor and Hierarchical Models: Specification, Inference, and Interpretation. *Annual Review of Clinical Psychology*, 15:51–69.

Marsh, H. W., Guo, J., Dicke, T., Parker, P. D., and Craven, R. G. (2019). Confirmatory Factor Analysis (CFA), Exploratory Structural Equation Modelling (ESEM), and Set-ESEM: Optimal Balance Between Goodness of Fit and Parsimony. *Multivariate Behavioral Research*, 26(4):529–556.

Marsh, H. W., Ludtke, O., Muthen, B. O., Asparouhov, T., Morin, A. J. S., and Trautwein, U. (2010). A New Look at the Big-Five Factor Structure through Exploratory Structural Equation Modeling. *Psychological Assessment*, 22:471–491.

Marsh, H. W., Ludtke, O., Nagengast, B., Morin, A. J. S., and Von Davider, M. (2013). Why Item Parcels are (almost) never appropriate: Two Wrongs do not Make a Right-Camouflaging Misspecification with Item Parcels in CFA models. *Psychological Methods*, 18:257–284.

Marsh, H. W., Morin, A. J. S., Parker, P. D., and Kaur, G. (2014). Exploratory Structural Equation Modeling: An Integration of the Best Features of Exploratory and Confirmatory Factor Analysis. *Annual Review of Clinical Psychology*, 10:85–110.

Marsh, H. W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. J. S., and Trautwein, U. (2009). Exploratory Structural Equation Modeling, Integrating CFA and EFA: Application to Students ' Evaluations of University Teaching. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(3):439–476.

McDonald, R. P. (1984). *Factor Analysis and Related Methods*. CRC Press, Boca Raton, FL.

Mcdonald, R. P. (1999). *Test Theory: A Unified Treatment*. Lawrence Erlbaum, Mahwah, N.J.

Meredith, W. (1977). On Weighted Procrustes and Hyperplane Fitting in Factor Analytic Rotation. *Psychometrika*, 42(4):491–522.

Millsap, R. E. (2001). When Trivial Constraints Are Not Trivial: The Choice of Uniqueness Constraints in Confirmatory Factor Analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, 8(1):1–17.

Molenaar, D. (2016). On the Distortion of Model Fit in Comparing the Bifactor Model and the Higher-Order Factor Model. *Intelligence*, 57:60–63.

Moore, T. M. (2013). *Iteration of Target Matrices In Exploratory Factor Analysis*. PhD thesis, University of California.

Moore, T. M., Calkins, M. E., Satterthwaite, T. D., Roalf, D. R., Rosen, A. F., Gur, R. C., and Gur, R. E. (2019). Development of a Computerized Adaptive Screening Tool for Overall Psychopathology. *Journal of Psychiatric Research*, 16:26–33.

Moore, T. M., Reise, S. P., Depaoli, S., and Haviland, M. G. (2015). Iteration of Partially Specified Target Matrices: Applications in Exploratory and Bayesian Confirmatory Factor Analysis. *Multivariate Behavioral Research*, 50(2):149–161.

Morin, A. J. S., Arens, A. K., and Marsh, H. W. (2016). A Bifactor Exploratory Structural Equation Modeling Framework for the Identifciation of Distinct Sources of Construct-Relevant Psychometric Multidimensionality. *Structural Equation Modeling: A Multidisciplinary Journal*, 23(1):116–139.

Mosier, C. I. (1939). Determining a Simple Structure when Loadings for Certain Tests are Known. *Psychometrika*, 4:149–162.

Mulaik, S. A. (1986). Factor Analysis and Psychometrika: Major Developments. *Psychometrika*, 51(1):23–33.

Mulaik, S. A. (2010). *Foundations of Factor Analysis. 2nd Edition*. CRC Press, New York, NY.

Mulaik, S. A. (2018). Identifying and Analyzing Scales Fundamentals of Common Factor Analysis. In Irwing, P., Booth, T., and Hughes, D. J., editors, *The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale and Test Development*. John Wiley & Sons, Ltd.

Murray, A. J. and Johnson, W. (2013). The Limitations of Model Fit in Comparing the Bi-Factor versus Higher-Order Models of Human Cognitive Ability Structure. *Intelligence*, 41(2013):407–422.

Myers, N. D., Ahn, S., and Jin, Y. (2013). Rotation to a Partially Specified Target Matrix in Exploratory Factor Analysis: How Many Targets? *Structural Equation Modeling: A Multidisciplinary Journal*, 20(1):131–147.

Myers, N. D., Ahn, S., Lu, M., Celimli, S., and Zopluoglu, C. (2017). Reordering and Reflecting Factors for Simulation Studies With Exploratory Factor Analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, 24(1):112–128.

Myers, N. D., Jin, Y., Ahn, S., Celimli, S., and Zopluoglu, C. (2015). Rotation to a partially specified target matrix in exploratory factor analysis in practice. *Behavior Research Methods*, 47(2):494–505.

Myers, N. D., Martin, J., Ntoumanis, N., Celimli, S., and Bartholomew, K. (2014). Exploratory Bi-factor Analysis in Sport, Exercise, and Performance Psychology: A Substantive Methodological Synergy. *Sport, Exercise, and Performance Psychology*, 3(4):258–272.

Nicewander, A. W. (2019). A Perspective of the Mathematical and Psychometric Aspects of Factor Indeterminacy. *Multivariate Behavioral Research*, pages 1–14.

Paunonen, S. V. (1997). On Chance and Factor Congruence Following Orthogonal Procrustes Rotation. *Educational and Psychological Measurement*, 57(1):33–59.

Peeters, C. F. W. (2012). *Bayesian Exploratory and Confirmatory Factor Analysis Perspectives on Constrained-Model Selection*. PhD thesis, Dept. of Methodology and Statistics, Utrecht University.

Reise, S. P. (2012). The Rediscovery of Bifactor Measurement Models. *Multivariate Behavioral Research*, 47(5):667–696.

Reise, S. P., Bonifay, W., and Haviland, M. G. (2013). Scoring and Modeling Psychological Measures in the Presence of Multidimensionality. *Journal of Personality Assessment*, 95(2):129–140.

Reise, S. P., Bonifay, W., and Haviland, M. G. (2018). Evaluating Scales Bifactor Modelling and the Evaluation of Scale Scores. In Irwing, P., Booth, T., and Hughes, D. J., editors, *The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale and Test Development*, chapter 22, pages 677–708. John Wiley & Sons, Ltd., first edition edition.

Reise, S. P., Kim, D. S., Mansolf, M., and Widaman, K. F. (2016). Is the Bifactor Model a Better Model or Is It Just Better at Modeling Implausible Responses? Application of Iteratively Reweighted Least Squares to the Rosember Self-Esteem Scale. *Multivariate Behavioral Research*, 51(1):818–838.

Reise, S. P., Moore, T. M., and Haviland, M. G. (2010). Bifactor Models and Rotations: Exploring the Extent to Which Multidimensional Data Yield Univocal Scale Scores. *Journal of Personality Assessment*, 92(6):544–559.

Reise, S. P., Moore, T. M., and Maydeu-Olivares, A. (2011). Target Rotations and Assessing the Impact of Model Violations on the Parameters of Unidimensional Item Response Theory Models. *Educational and Psychological Measurement*, 71(4):684–711.

Reise, S. P., Morizot, J., and Hays, R. D. (2007). The role of the Bifactor Model in Resolving Dimensionality Issues in Health Outcomes Measures. *Quality of Life Research*, 16:19–31.

Reise, S. P., Scheines, R., Widaman, K. F., and Haviland, M. G. (2012). Multidimensionality and Structural Coefficient Bias in Structural Equation Modeling: A Bifactor Perspective. *Educational and Psychological Measurement*, 73(1):5–26.

Revelle, W. and Condon, D. M. (2019). Reliability from alpha to omega: A tutorial. *Psychological Assessment*, 31(12):1395–1411.

Revelle, W. and Wilt, J. (2013). The general factor of personality: A general critique. *Journal of Research in Personality*, 47(5):493–504.

Rigdon, E. E., Becker, J.-M., and Sarstedt, M. (2019). Factor Indeterminacy as Metrological Uncertainty: Implications for Advancing Psychological Measurement. *Multivariate Behavioral Research*, 54(3):429–443.

Robertson, S. (2019). *Bifactor Models and Factor Collapse: A Monte Carlo Study*. PhD thesis, Clemson University.

Rodriguez, A., Reise, S. P., and Haviland, M. G. (2015). Evaluating Bifactor Models: Calculating and Interpreting Statistical Indices. *Psychological Methods*, 21(2):137–150.

Rodriguez, A., Reise, S. P., and Haviland, M. G. (2016). Applying Bifactor Statistical Indices in the Evaluation of Psychological Measures. *Journal of Personality Assessment*, 98(3):223–237.

Ronald, A. (2019). Editorial: The Psychopathology P Factor: Will it Revolutionise the Science and Practice of Child and Adolescent Psychiatry. *Journal of Child Psychology and Psychiatry*, 60(5):497–499.

Rozeboom, W. W. (1991). Hyball: A Method for Subspace-Constrained Factor Rotation. *Multivariate Behavioral Research*, 26(1):163–177.

Rozeboom, W. W. (1992). The Glory of Suboptimal Factor Rotation: Why Local Minima in Analytic Optimization of Simple Structure are More Blessing Than Curse. *Multivariate Behavioral Research*, 27(4):585–599.

Scharf, F. and Nestler, S. (2019a). A Comparison of Simple Structure Rotation Criteria in Temporal Exploratory Factor Analysis for Event-Related Data Potential. *Methodology*, 15:43–60.

Scharf, F. and Nestler, S. (2019b). Should Regularization Replace Simple Structuer Rotation in Exploratory Factor Analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, 26(4):576–590.

Schönemann, P. H. (1966). A Generalized Solution of the Orthogonal Procrustes Problem. *Psychometrika*, 31(1):1–10.

Sellbom, M. and Tellegen, A. (2019). Factor Analysis in Psychological Assessment Research: Common Pitfalls and Recommendations. *Psychological Assessment*, pages 1–14.

Sočan, G. (2003). *The Incremental Value of Minimum Rank Factor Analysis*. PhD thesis, University of Groningen.

Spearman, C. (1904). 'General Intellligence', Objectively Determined and Measured. *The American Journal of Psychology*, 15(2):201–293.

Stegeman, A. (2016). A new Method for Simultaneous Estimation of the Factor Model Parameters, Factor Scores, and Unique Parts. *Computational Statistics and Data Analysis*, 99:189–203.

Steiger, J. H. (1996). Coming Full Circle in the History of Factor Indeterminacy. *Multivariate Behavioral Research*, 31(4):617–630.

Steiger, J. H. (2002). When Constraints Interact: A Caution About Reference Variables, Identification Constraints, and Scale Dependencies in Structural Equation Modeling. *Psychological Methods*, 7(2):210–227.

Swineford, F. (1941). Some Comparisons of the Multiple-Factor and the Bi-factor Methods of Analysis. *Psychometrika*, 6(6):375–382.

ten Berge, J. M. F. (2006). The Rigid Orthogonal Procrustes Rotation Problem. *Psychometrika*, 71(1):201–205.

ten Berge, J. M. F. and Nevels, K. (1977). A General Solution to Mosier's Oblique Procrustes Problem. *Psychometrika*, 42(4):593–600.

Thomson, G. H. (1916). A Hierarchy Without a General Factor. *British Journal of Psycohlogy*, 8(3):271–281.

Thurstone, L. L. (1933). *The Theory of Multiple Factors*. Edwards Brothers, Ann Arbor, Michigan.

Thurstone, L. L. (1940). Current Issues in Factor Analysis. *Psychological Bulletin*, 37(4):189–235.

Thurstone, L. L. (1947). *Multiple-Factor Analysis: A Development and Expansion of the Vectors of Mind*. Univ. Chicago Press, Chicago, IL, US.

Trendafilov, N. T. (1994). A Simple Method for Procrustean Rotation in Factor Analysis Using Majorization Theory. *Multivariate Behavioral Research*, 29(4):385–408.

Tucker, L. R. (1940). A Rotational Method Based upon the Mean Principal Axis of a Subgroup of Tests. *Psychological Bulletin*, 37:578.

Tucker, L. R. (1944). A Semi-Analytical Method of Factor Rotation to Simple Structure. *Psychometrika*, 9(1):43–68.

van Bork, R., Epskamp, S., Rhemtulla, M., Borsboom, D., and van der Maas, H. L. J. (2017). What is the P-factor of Psychopathology? Some Risks of General Factor Modeling. *Theory and Psychology*, 27(6):759–773.

van der Maas, H. L. J., Dolan, C. V., Grasman, R. P., Wicherts, J. M., Huizenga, H. M., and Raijmakers, M. E. (2006). A Dynamical Model of General Intelligence: The Positive Manifold of Intelligence by Mutualism. *Psychological Review*, 13:842–860.

Waller, N. G. (2017). Direct Schmid–Leiman Transformations and Rank-Deficient Loadings Matrices. *Psychometrika*, pages 1–13.

Waller, N. G. (2018). An Introduction to Kristof's Theorem for Solving Least-Square Optimization Problems Without Calculus. *Multivariate Behavioral Research*, 53(2):190–198.

Watt, A. L., Poore, H. E., and Waldman, I. D. (2019). Riskier Test of the Validity of the Bifactor Model of Psychopatology. *Clinical Psychological Science*, 7(6):1285–1303.

Weide, A. C. and Beauducel, A. (2019). Varimax Rotation Based on Gradient Projecton is a Feasible Alternative to SPSS. *Frontiers in Psychology*, 10:1–15.

Wherry, R. J. (1959). Hierarchical Factor Solutions Without Rotation. *Psychometrika*, 24(1):54–51.

Wilson, E. B. (1928). Coming Full Circle in the History of Factor Indeterminacy. *Science*, 67(1731):244–248.

Xiao, Y., Liu, H., and Hau, K. T. (2019). A Comparison of CFA, ESEM, and BSEM in Test Structure Analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, 26(5):665–677.

Yamamoto, M., Hirose, K., and Nagata, H. (2017). Graphical Tool of Sparrse Factor Analysis. *Biometrika*, 44(1):229–250.

Yamashita, N. and Adachi, K. (2019). Permutimin: Factor Rotation to Simple Structure with Permutation of Variables. *Multivariate Behavioral Research*, 55(1):17–29.

Yang, R., Spirtes, P., Scheines, R., Reise, S. P., Yang, R., Spirtes, P., Scheines, R., Reise, S. P., and Mansoff, M. (2017). Finding Pure Submodels for Improved Differentiation of Bifactor and Second-Order Models. *Structural Equation Modeling: A Multidisciplinary Journal*, 24(3):402–413.

Yates, A. (1987). *Multivariate Exploratory Data Analysis: A Perspective on Exploratory Factor Analysis*. State University of New York Press, Albay, N.Y.

Yung, Y.-F., Thissen, D., and McLeod, L. D. (1999). On the Relationship Between The Higher-Order Factor Model and The Hierarchical Factor Model. *Psychometrika*, 64(2):113–128.

Zhang, B., Sun, T., Cao, M., and Drasgow, F. (2020). Using Bifactor Models to Examine the Predictive Validity of Hierarchical Constructs: Pros, Cons, and Solutions. *Organizational Research Methods*.

Zhang, G., Hattori, M., Trichtinger, L. A., and Wang, X. (2018a). Target Rotation with Both Factor Loadings and Factor Correlations. *Psychological Methods*, 24(3):390–402.

Zhang, G., Preacher, K. J., Hattori, M., Jiang, G., and Trichtinger, L. A. (2018b). A Sandwich Standard Error Estimator for Exploratory Factor Analysis With Nonnormal Data and Imperfect Models. *Applied Psychological Measurement*, 43(5):360–373.

Zinbarg, R. E. and Alden, A. R. (2015). Coefficient Alpha and Coefficient Omega Hierarchical. *The Encyclopedia of Clinical Psychology*, pages 1–6.

Zinbarg, R. E., Revelle, W., and Yovel, I. (2007). Estimating ω_H for Structures Containing two Group Factors: Perils and Prospects. *Applied Psychological Measurement*, 31(2):135–157.

Zinbarg, R. E., Revelle, W., Yovel, I., and Li, W. (2005). Cronbach's, α Revelle's β and McDonald's ωH: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, 70(1):123–133.

Appendix A

Contributed Work

A.1 Main Work

1. **Garcia-Garzon, E.**, Nieto, M.D., Abad, F.J. & Garrido, L.E. (submitted). Bi-factor ESEN Done Right: Using the SLiDapp Application. Psicothema. (IF = 1.551).

2. **Garcia-Garzon, E.**, Abad, F.J. & Garrido, L.E. (in press). On Omega Hierarchical Estimation: A Comparison of Exploratory Bi-factor Analysis Algorithms. Multivariate Behavioral Research. (IF = 2.141).

3. **Garcia-Garzon, E.**, Abad, F.J. & Garrido, L.E. (2019). Searching for G: A New Evaluation of SPM-LS dimensionality. The Journal of Intelligence, 7 (3), 14. doi: 10.3390/jintelligence7030014

4. **Garcia-Garzon, E.**, Abad, F.J. & Garrido, L.E. (2019). Improving Bi-factor Exploratory Modelling: Empirical Target Rotation Based on Loading Differences. Methodology, 15 (2), 45-55, doi: 10.1027/1614-2241/a000163 (IF = 0.900).

5. Abad, F.J., **Garcia-Garzon, E.**, Garrido, L.E. & Barrada, J.R. (2017). Iteration of Partially Specified Target Matrices: Application to the Bi-Factor Case. Multivariate Behavioral Research, 52 (4), 416-429. doi: 10.1080/00273171.2017.1301244 (IF = 3.691).

A.2 Conference: Oral Presentations

1. **Garcia-Garzon, E.**, Abad, F.J. & Garrido, L.E. (July 2019). Assessing general factor reliability in exploratory bi-factor modelling. The 2019 International Meeting of the

Psychometric Society, Pontificia Universidad Católica de Chile, Santiago de Chile, Chile.

2. **Garcia-Garzon, E.** & Lozano, A. M. (July 2019). Hacia un análisis de datos exploratorio transparente. The XVI Congreso de Metodología de las Ciencias Sociales y de la Salud, Universidad Autónoma de Madrid, Madrid, España.

3. **Garcia-Garzon, E.** & Lozano, A. M. (July 2019). Sesgos de Respuesta en Modelos de Redes Neuronales Artificiales: Aplicaciones al Análisis de Relaciones No Lineales. The XVI Congreso de Metodología de las Ciencias Sociales y de la Salud, Universidad Autónoma de Madrid, Madrid, España.

4. **Garcia-Garzon, E.**, Abad, F.J. & Garrido, L.E. (June 2018). Improving Bi-factor Exploratory Modelling: Empirical Target Rotation Based on Loading Differences. The 2018 International Meeting of the Psychometric Society, University of Columbia, New York, United States.

5. **Garcia-Garzon, E.**, Abad, F.J. & Garrido, L.E. (June 2018). Exploratory Bi-factor Model-Base Reliability: The Omega Hierarchical Case. The VII European Congress of Methodology, University of Jena, Jena, Germany.

6. **Garcia-Garzon, E.**, Abad, F.J. & Garrido, L.E. (July 2017). Weak Factor Recovery in Bi-factor Exploratory Analysis: A Comparison of Empirical versus Arbitrary Cut-off Selection Criteria. XV Congreso de Metodología de las Ciencias Sociales y de la Salud, Barcelona, España.

A.3 Conference: Poster Presentations

1. **Garcia-Garzon, E.**, Abad, F.J. & Garrido, L.E. (July 2017). Empirical versus Arbitrary Cut-off Points in Exploratory Bi-factor Target Rotation. Poster presented at the 2017 International Meeting of the Psychometric Society, Zurich, Switzerland

A.4 Research Awards & Scholarships

1. 2016-2020: 4-year FPU Fellowship (FPU 15/03246) from Spanish Ministry of Education, Culture and Sport, Spain. (75000 euro).

2. 2019: The Journal of Intelligence Travel Award (1000 swiss franc).

3. 2019: Julio Olea Prize for Young Researchers in Methodology. Asociación Española de Metodología en Ciencias del Comportamiento y la Salud. Prize (2000 euro).

4. 2018: FPU Fellowship Visiting Scholar Grant at University of Groningen, the Netherlands (September 2018 – December 2018). (3200 euro)

5. 2018: Best Presentation at Annual Doctorate Seminar, Universidad Autónoma de Madrid.

6. 2017: Best Junior Researcher Presentation at XV Congreso de la Asociación Española de Metodología en Ciencias del Comportamiento y la Salud, Barcelona (1000e)

A.5 Participation in Research Funded Projects

1. 2019-2020- PSI2013-44300-P, Spanish Ministry of Economy, Industry and Competitiveness.

2. 2017-2019- PSI2017-85022-P, Spanish Ministry of Economy, Industry and Competitiveness.

3. 2019-2020- PS-001.19-INN, Universidad Autónoma de Madrid.

Appendix B

Resumen

El modelo bi-factorial se ha convertido en uno de los principales modelos estadísticos dentro de la investigación en psicología. Gracias a sus características únicas, el modelo bi-factorial ha sido redescubierto en las principales áreas de investigación en psicología: el estudio de la psicopatología, personalidad o inteligencia. Así, nuestro conocimiento sobre este modelo puede determinar futuros avances clave en la investigación de los fenómenos psicológicos mencionados. Pese a los importantes esfuerzos realizados para estimar estos modelos bi-factoriales de una manera fiable, los métodos de los que disponemos en la actualidad presentan severas limitaciones que ponen en cuestión su validez y utilidad.

Esta tesis doctoral revisará tanto el desarrollo matemático e histórico del modelo bi-factorial como sus principales aplicaciones desde perspectivas confirmatorias y exploratorias. Así, la tesis se centrará en el modelo bi-factorial exploratorio al considerarse en la actualidad la opción más adecuada para aproximar dicho modelo en las condiciones más habituales de su aplicación. Adicionalmente, la tesis buscará aclarar el rol central que ha desempeñado la rotación target en la aproximación de los modelos bi-factoriales exploratorios, destacando las limitaciones y potenciales áreas de mejoras de los métodos actuales. Esta tesis doctoral tiene como objetivo desarrollar nuevos algoritmos de rotación bi-factorial que permitan mejorar la estimación de parámetros y otros estadísticos de interés en estos modelos (fiabilidad, etc.). Adicionalmente, se buscará ofrecer herramientas de libre acceso para facilitar el uso de los métodos desarrollados y la aproximación al análisis bi-factorial exploratorio por parte de investigadores especializados. En detalle, estos objetivos se desarrollaron en siete capítulos.

En el Capítulo 2 se introdujo la Rotación Target Iterativa basada en una solución de Schmid-Leiman (i.e., SLi). Este algoritmo mejoraba la propuesta original de Reise, Moore and Maydeu-Olviares (2011) de definir una matriz de rotación target basada en Schmid-Leiman mediante la aplicación del algoritmo de rotación iterativa de Moore, Reise, Depaoli and Haviland (2016). Los resultados de un estudio Monte Carlo demostraron que SLi resultaba en una mejor recuperación de la estructura bi-factorial cuando se comparaba con otros métodos tales como bi-geomin, bi-quartimin, la solución original de Schmid-Leiman o la rotación target basada en Schmid-Leiman no iterativa.

En el Capítulo 3 se presentó la Rotación Target Empírica Iterativa basada en una solución de Schmid-Leiman (i.e., SLiD). Este algoritmo actualizaba SLi mediante la introducción de un nuevo método para calcular puntos de corte específicos para cada factor en la definición de la matriz target. Esta estrategia se basaba en encontrar saltos relevantes en la distribución de diferencias de los pesos factoriales normalizados y ordenados de cada factor de grupo. Los resultados de un estudio Monte Carlo demostraron la superioridad del algoritmo SLiD frente a SLi en condiciones realistas: estructuras bi-factoriales que incluían una mezcla de factores de grupo con diferentes pesos medios y que presentaban un número alto de pesos cruzados.

En el Capítulo 4 se aplicó la estrategia subyacente al algoritmo SLiD al estudio de las puntuaciones de un instrumento breve para medir inteligencia: las últimas doce matrices de las Matrices Progresivas del Raven (SPM-LS). Esta aplicación demostró la utilidad de ese enfoque para evaluar tanto el supuesto de unidimensionalidad así como la presencia de factores menores. Gracias a la aplicación de los métodos desarrollados previamente, se demostró que el modelo bi-factorial representaba el modelo de medida más adecuado en dicho contexto.

En el Capítulo 5 se exploraron las consecuencias que tenía la elección de un algoritmo bi-factorial exploratorio en la estimación del estadístico omega jerárquico. Además, en este estudio se incluyeron dos nuevos algoritmos: Direct Bi-factor y Direct Schmid-Leiman. De nuevo, los resultados de tres simulaciones Monte Carlo evidenciaron que el algoritmo SLiD favoreció una mejor estimación de omega jerárquico a través de una multitud de estructuras (i.e., un modelo bi-factorial de rango completo o deficiente, o estructuras sin factor general). Adicionalmente, se demostró que la elección del tipo de rotación target (parcial o totalmente especificada) determinaba la calidad de la estimación de omega jerárquico en los algoritmos basados en la rotación target. Por último, se ejemplificó el funcionamiento de todos los algoritmos en ocho ejemplos empíricos clásicos de la literatura del área.

En el Capítulo 6 se integró el algoritmo SLiD dentro del contexto de los modelos de ecuaciones estructurales exploratorias. Sin embargo, como este algoritmo no se encuentra todavía disponible en el principal software para realizar estos análisis (Mplus), se desarrolló una aplicación web Shiny libre y fácil de usar (*SLiDApp*) para que los usuarios trasladasen la matriz de rotación estimada por SLiD a Mplus. Para facilitar su uso, se desarrolló una guía paso a paso con indicaciones de uso de esta aplicación que incluía un análisis de ecuaciones exploratorias bi-factoriales para estudiar la relación entre creencias en teorías conspirativas y rasgos de personalidad.

Esta tesis doctoral concluye con una discusión crítica y reflexiva de los méritos e inconvenientes del modelo bi-factorial exploratorio y de los métodos existentes para su estimación. De este modo se buscan enfatizar tanto las futuras líneas de investigación como consejos para aplicar estos modelos dirigidos a la comunidad de investigadores en psicología. Por ello se espera que la tesis doctoral sea útil para desentrañar las ventajas y las limitaciones actuales de los modelos bi-factoriales, así como que sirva de inspiración para futuras investigaciones.

Appendix C

Discusión General

C.1 Resultados Principales

El modelo bi-factorial se ha convertido en uno de los principales modelos estadísticos dentro de la investigación en psicología (Markon, 2019). Pese a ello, el modelo bi-factorial es todavía pobremente entendido (Bonifay et al., 2017; Markon, 2019). Así, sigue siendo comúnmente aplicado desde una perspectiva confirmatoria que resulta inadecuada en la mayoría de contextos de investigación (Reise et al., 2018, 2011). Esta tesis doctoral ha sido dedicada a resolver las dificultades asociadas a la aproximación exploratoria de estos modelos. Para ello, se realizó una revisión y discusión detallada de su historia y su asociación con los conceptos de estructura simple y rotación factorial. Adicionalmente, se describió el rol prominente de la rotación target en la aproximación del modelo bi-factorial exploratorio (Browne, 2001; Fleming, 2012). Ya sea por su simplicidad matemática así como su habilidad para recuperar estructuras factoriales complejas, la rotación target está posicionada para seguir desempeñando un papel importante en futuras aproximaciones a modelos factoriales complejos dentro y fuera de caso bi-factorial (Guo et al., 2019; Marsh et al., 2019). Por lo tanto, esta tesis se centró en desarrollar nuevos métodos que permitan perfeccionar la aproximación de los modelos bi-factoriales exploratorios usando este tipo de rotación. Esta tesis doctoral ha buscado que mediante la presentación de los avances conseguidos dentro de un marco unificado (la familia SLi de rotaciones) se haya favorecido la posibilidad de que existan colaboraciones futuras para su desarrollo[1].

En detalle, los Capítulos 2 y 3 de la tesis fueron dedicados a la introducir los algoritmos SLi y SLiD como herramientas clave para la estimación de modelos bi-factoriales explorato-

[1]Un primer ejemplo fue la extensiva correspondencia con el Dr. Stephen Robertson durante sus estudios relativos al colapso factorial en modelos bi-factoriales. Desde esta nota, agradecerle personalmente por compartir su conocimiento y entusiasmo por el tema durante estos años.

rios. El Capítulo 4 demostró su aplicación a la evaluación de la estructura factorial del test de las doce ultimas matrices de las matrices progresivas de Raven (i.e., SPM-LS). Además, dada la relevancia de omega jerárquico como estadístico de referencia en el área, el Capítulo 5 introdujo una detallada evaluación de cómo la elección del algoritmo de rotación bi-factorial podría afectar la calidad de su estimación. Por último, en el Capítulo 6 se evaluaron los diferentes métodos de rotación dentro del contexto de los modelos ESEM. Para ello, se desarrolló y ejemplificó el uso de SLiDApp, una aplicación que permitía trasladar de manera sencilla la rotación bi-factorial calculada en SLiD de R a *Mplus*para su uso en este tipo de modelos estructurales.

Para concluir, cabe destacar que esta tesis doctoral ha intentado clarificar algunos de los puntos más controvertidos de la literatura relativa a los modelos bi-factoriales. Además de proveer una comparación detallada de los diferentes algoritmos disponibles para estimar los modelos bi-factoriales de manera exploratoria, esta tesis ha generado otros resultados de interés que no deberían ser pasados por alto: nuevas taxonomías y conceptualizaciones de los modelos bi-factoriales, aproximaciones alternativas para medir la recuperación de estos modelos y una nueva estrategia para la definición de la rotación target parcialmente especificada. Para clarificar estos puntos, los resultados principales de cada capítulo son detallados a continuación.

C.1.1 Capítulo 2: Principales Resultados

En el Capítulo 2 se ilustró cómo la aplicación iterativa de la rotación target parcialmente especificada mediante el algoritmo SLi tenía un efecto beneficioso en la recuperación de los modelos bi-factoriales exploratorios. Esta conclusión se basó en los resultados obtenidos tanto en una simulación Monte Carlo como en el análisis de un ejemplo empírico. La simulación presentaba un extenso conjunto de condiciones, incluyendo manipulaciones del tamaño muestral, el número de factores de grupo, el número de indicadores por factor y la simplicidad de la estructura bi-factorial. Particularmente, la simplicidad se manipuló mediante la simulación de pesos cruzados e ítems sin carga factorial en el factor de grupo dentro de la estructura factorial. Los resultados mostraron que SLi presentaba un mejor rendimiento que otros métodos en las condiciones más realistas de aplicación: cuando los factores de grupos incluían pesos cruzados o, de manera más relevante, ítems puros del factor general. Por ello, se recomendó su uso general para estimar estructuras bi-factoriales exploratorias. En este Capítulo se estudiaron además otros algoritmos de rotación bi-factorial: bi-quartimin, bi-geomin, SL y la rotación target parcialmente especificada basada en SL no iterativa. Los resultados mostraron que bi-quartimin no resultaba en una recuperación ante la presencia de pesos cruzados, que SL nunca recuperaba correctamente estructuras incluyendo indicadores puros del factor general y que bi-geomin mostraba la mejor recuperación de todos los métodos para las estructuras con un mayor grado de complejidad. Por último, se observó que ningún método proveía una buena recuperación factorial bajo tamaño muestrales pequeños (i.e., N = 200) o si los factores de grupos tenían pesos factoriales medios de magnitud baja (i.e., entre 0,30 y 0,50).

Esta investigación fue uno de los primeros estudios que realizó una comparación directa del funcionamiento de diferentes algoritmos de rotación bi-factorial usando métodos Monte Carlo. Así, no solo ha sido citado más de trece veces en la literatura (febrero, 2020), sino que las condiciones aquí estudiadas han sido usadas como base de otros muchos estudios en el área (Giordano and Waller, 2019; Lorenzo-Seva and Ferrando, 2019b; Robertson, 2019). Adicionalmente, este estudio introdujo una terminología nueva en el campo, inspirada en Mcdonald (1999) para distinguir la complejidad de una estructura bi-factorial cuando existen ítems puros: estructuras de clústeres independientes con indicadores puros (i.e., ICP) y estructuras de base de clústeres independientes con ítems puros (i.e., ICBP).

C.1.2 Capítulo 3: Resultados Principales

El Capítulo 3 presentó el algoritmo de rotación target empírica e iterativa basada en una solución de Schmid-Leiman (i.e., SLiD). Así, este capítulo incluía una descripción del nuevo método desarrollado para establecer puntos de corte empíricos basados en encontrar saltos en la distribución de diferencias entre pesos factoriales consecutivos para cada factor de grupo. Asimismo, y a diferencia de otros métodos similares, este Capítulo detallaba como esta estrategia se basaba en evitar fijar incorrectamente un peso cruzado relevante a cero frente a la maximización directa del número total de elementos a fijar en la matriz target.

El algoritmo SLiD se comparó con el algoritmo SLi que usaba diferentes puntos de cortes únicos para definir la matriz target (i.e., desde 0,05 a 0,20 a saltos de magnitud de 0,05). Se diseñó una simulación Monte Carlo para comparar estos algoritmos. Las estructuras simuladas buscaban reproducir condiciones realistas de investigación donde: (a) se incluían un número extensivo de pesos cruzados simulados para cada factor de grupo en base una distribución normal con tres amplitudes diferentes; y (b) existía una mezcla de factores de grupos fuertes, medios y débiles (i.e., factores de grupos que difieren en la magnitud de peso factorial medio). Adicionalmente, otras variables estudiadas en el Capítulo 2 fueron tomadas en consideración en esta simulación. El análisis de estos algoritmos concluyó con un reanálisis del ejemplo empírico incluido en ese estudio.

Los resultados reflejaron que el algoritmo SLiD resultaba en una mejor recuperación factorial que el algoritmo SLi (independientemente del punto de corte fijo que se usase). Particularmente, SLiD mostró una correcta recuperación de los factores de grupo con un alto número de pesos cruzados. En paralelo, se demostró que el algoritmo SLiD resultaba, en un número pequeño de condiciones, en colapso factorial y recuperación inadecuada de la estructura. Este resultado era consistente con la simulación, donde las estructuras incluían en cada factor de grupo un número alto de pesos cruzados de pequeña magnitud. Así, la adecuación del uso de SLi o un punto de corte determinado parecía depender de las condiciones de aplicación. Al final, se recomendaba a los investigadores que evitasen tomar decisiones inadecuadas y utilizasen el algoritmo SLiD, ya que mostraba tener un mejor comportamiento estadístico en una mayoría de condiciones, revisando siempre la calidad de las soluciones estimadas. Por último, se observó que los factores de grupo con un peso medio de baja magnitud eran difícilmente recuperables, particularmente si presentaban pesos cruzados de mayor magnitud.

C.1.3 Capítulo 4: Resultados Principales

El Capítulo 4 buscaba ilustrar a utilidad de los métodos diseñados en los Capítulos 2 y 3 en el contexto de la validación de una escala breve de inteligencia, las doce ultimas matrices de las Matrices Progresivas del Raven (i.e., SPM-LS). En este caso, se examinó al detalle el supuesto de unidimensionalidad de esta escala, crucial para las técnicas estadísticas aplicadas en el contexto original de la validación. Aunque tanto los métodos modernos como tradicionales de medición de la dimensionalidad (i.e., análisis paralelo y análisis gráfico exploratorio) revelaron fuentes de variabilidad adicionales al factor general, los resultados del bi-factorial apoyaron que las puntuaciones del SPM-LS se pudieran considerar esencialmente unidimensionales. Sin embargo, se identificó un factor menor correspondiente a los seis últimos ítems del test. Así, aunque las puntuaciones totales del SPM-LS puedan ser usadas como medidas válidas de diferencias individuales en inteligencia general, el modelo correcto de medida de carácter bi-factorial (que incluía dicho factor menor) debería ser considerado en aquellas situaciones donde, por ejemplo, el ajuste a los datos deba prevalecer frente a otros criterios.

C.1.4 Capítulo 5: Resultados Principales

El Capítulo 5 se dedicó a realizar una comparación detallada de los algoritmos de rotación bi-factorial con respecto a la recuperación del estadístico de omega jerárquico. En este estudio se incluyeron, por primera vez, dos nuevos algoritmos a la comparativa: Direct Schmid-Leiman y Direct Bi-factor (Giordano and Waller, 2019; Waller, 2017). Todos los métodos fueron comparados en tres simulaciones Monte Carlo investigando la recuperación de omega jerárquico para modelos bi-factoriales de rango completo, modelos jerárquicos transformados mediante la ortogonalización de SL y una simulación final incluyendo estructuras sin un factor general. Adicionalmente, la recuperación de omega se estudió mediante el error medio absoluto (i.e., MAE), el sesgo medio (i.e., MBE) y la raíz cuadrada del error medio cuadrático (i.e., RMSE). Por último, y a diferencia de lo que ocurría en el Capítulo 2, se modificó la magnitud del peso medio del factor general para estudiar su efecto en la recuperación de omega jerárquico.

Los resultados mostraron que el algoritmo SLiD proveía la mejor estimación de omega jerárquico en una mayoría de condiciones y estructuras. Sin embargo, ni siquiera SLiD pudo mostrar una buena recuperación de este estadístico bajo muestras pequeñas (i.e., N = 150), un número pequeño de factores (i.e.., 3) o si el factor general tenía un peso factorial medio de baja magnitud (i.e., 0,30). Consecuentemente, se recomendó a los investigadores que desconfiasen de cualquier estimación de omega jerárquico bajo estas condiciones. SL, bi-quartimin y bi-geomin no recuperaron omega jerárquico de manera adecuada, ya que tendían a sobreestimar su valor. Mientras que SL resultó infraestimó omega jerárquico, bi-geomin y bi-quartimin sobreestimaron dicho estadístico. En este sentido, la infraestimación de SL se atribuyó, por primera vez, a una estimación incorrecta de la matriz de correlaciones de primer orden sobre la que extrae el factor general. Por último, el funcionamiento de los algoritmos se evaluó usando ocho ejemplos clásicos de la literatura bi-factorial y de los modelos de segundo orden.

El estudio de los métodos Direct Schmid-Leiman y Direct Bi-factor reveló algunos aspectos interesantes de cómo funcionan los métodos basados en una rotación target completamente especificada: la calidad de la rotación depende de la existencia de discrepancias entre la media de los pesos factoriales de los factores de la estructura bi-factorial. Cuando los factores tienen pesos factoriales homogéneos, una rotación target completamente especificada resultará en una estimación insesgada de los pesos de la misma, y, por ende, de omega jerárquico. Por otro lado, en el momento en que factores difieran en su peso factorial medio, los pesos rotados estarán sesgados hacia la media de los pesos que han sido especificados dados un valor target de uno (i.e., han sido maximizados). Además, se encontró que este efecto tenía graves consecuencias en la estimación de omega jerárquico, ya que las pequeñas y sistemáticas desviaciones provocadas en el factor general por estos métodos tenían un alto impacto en el valor de omega jerárquico (independientemente de la calidad de la recuperación general de los pesos factoriales). Los resultados informaron que la relación entre estimar correctamente los pesos factoriales y el sesgo en la estimación de estadísticos secundarios calculados en función de estas soluciones no es tan simple como pudiera parecer. La presencia de desviaciones sistemáticas es más relevante para el segundo caso que para el primero. Así, asegurar una buena recuperación de la estructura factorial podría ser una condición necesaria, pero no suficiente, para la correcta estimación de este tipo de estadísticos. Dicho resultado novedoso representa un avance importante no contemplado hasta ahora de manera explícita en la literatura relativa a la comparación de métodos target completa y parcialmente especificados. Por último, el funcionamiento de estos algoritmos se ejemplificó con el análisis de ocho datasets clásicos de la literatura bi-factorial.

C.1.5 Capítulo 6: Resultados Principales

El Capítulo 6 se dedicó a ilustrar la utilidad del algoritmo SLiD en el contexto de los modelos de ecuaciones estructurales exploratorias (i.e., ESEM). ESEM permite expandir los modelos de ecuaciones estructurales tradicionales para acomodar modelos de medición exploratorios (Marsh et al., 2014, 2009), y que han demostrado superar a otras alternativas tradicionales en multitud de condiciones (Asparouhov and Muthén, 2009; Guo et al., 2019; Marsh et al., 2019). Así, los modelos ESEM bi-factoriales han atraído una importante atención en la literatura (Morin et al., 2016). Desafortunadamente, el principal software para estimar ESEM (i.e., *Mplus*) únicamente ofrece bi-geomin, bi-quartimin o la rotación target no iterativa como métodos disponibles para aproximar estructuras bi-factoriales exploratorias. En este capítulo se ilustró como llevar a cabo un análisis ESEM bi-factorial usando SLiD gracias a una aplicación Shiny (i.e., SLiDApp) desarrollada para tal efecto. Esta aplicación permitía estimar una matriz target basada en SLiD y trasladarla para su uso en un modelo ESEM en *Mplus*. Este proceso se ejemplificó paso a paso mediante un análisis ESEM bi-factorial de la relación entre las puntuaciones en el Test General de Creencias Conspirativas y los cinco factores de personalidad. Los resultados mostraron que, como se esperaba, SLiD permitió recuperar aspectos únicos de los datos cuando en comparación con otros algoritmos. Asimismo, se esperó que gracias a la guía ofrecida para realizar ESEM bi-factorial basado en SLiD, futuros investigadores interesados en modelos bi-factoriales ESEM se beneficien del uso de la aplicación Shiny diseñada.

C.2 Futuras Direcciones y Limitaciones

C.2.1 La Naturaleza del Modelo bi-factorial

Es necesario comenzar esta sección enfatizando que, en la actualidad, todavía desconocemos una gran parte de las propiedades estadísticas del modelo bi-factorial. Pese a que han pasado de una década desde su "redescubrimiento" (Reise et al., 2012), existen hoy todavía un número importante de preguntas sin resolver relativas a este modelo (Bonifay et al., 2017; Markon, 2019). Y la primera y más importante de todas ellas es...¿qué es un modelo bi-factorial?

Desafortunadamente, todavía no podemos responder a dicha pregunta con seguridad. En una mayoría de ocasiones, el modelo bi-factorial es una etiqueta que se da a toda matriz de pesos factoriales que incluye un factor general y varios factores de grupos (que se espera que conformen un patrón de estructura simple). Así, artículos recientes incluyen las soluciones SL como una subclase de modelos bi-factoriales de rango deficiente (Waller, 2018; Giordiano & Waller, 2019). Sin embargo, y como se explica en el Capítulo 1, esta decisión puede ser, cuanto menos, cuestionable: no todas las estructuras bi-factor de rango deficientes son consistentes con una transformación de tipo SL. Por lo tanto, merecería la pena distinguir entre soluciones de rango deficiente que son consistentes con un modelo generativo de una solución de segundo orden (independientemente de si son obtenidas mediante la ortogonalización de SL u otro método) y soluciones de rango deficiente que no son consistentes con dicho modelo jerárquico. Además, cabe recalcar que la literatura se beneficiaría de una mayor distinción entre el modelo de segundo orden y el uso de la transformación de SL como un método para convertir las soluciones en estructuras "bi-factoriales". Como decía Gignac (2016): "El modelo jerárquico es un modelo que, por ejemplo, puede ser especificado y cuya plausibilidad puede ser comprobada estadísticamente. La transformación de Schmid-Leiman, sin embargo, no puede ser especificada ni su plausibilidad, comprobada. Simplemente es un método usado para calcular efectos indirectos" (p.58, nota al pie 1).

Otro debate relacionado es si los modelos bi-factoriales forman parte de la familia de modelos jerárquicos o no (Markon, 2019). Cómo Yung et al. (1999) demostraron, es útil encajar ambos tipos de modelos dentro de un marco unificado para entender sus similitudes y diferencias, particularmente cuando los investigadores están interesados en evaluar las restricciones estadísticas implícitas en cada uno. Así, muchos autores han optado por referirse al modelo jerárquico y al bi-factorial como modelos jerárquicos indirectos y directos, respectivamente (Gignac, 2008, 2016). Esta distinción pone el acento en las implicaciones teóricas de cada modelo y en la presencia de los efectos directos o mediados del factor general a los ítems en cada caso. Sin embargo, no existe un acuerdo unánime sobre el uso de dicha taxonomía: "[sobre el modelo bi-factorial] Es necesario destacar que no es un modelo jerárquico, ya que g (que necesariamente pesa en todos los ítems) no depende de los pesos de las variables en los factores de grupos" (Jennsen & Weng, 1994, p.245). Este debate, aunque técnico, no debe ser tomado a la ligera. La consideración del modelo bi-factorial como parte de la familia de modelos jerárquicos puede afectar a su consideración como un modelo de medida aceptable (i.e., consistente con la teoría) en ciertas áreas de investigación (i.e., inteligencia o psicopatología). En cualquier caso, lo que está fuera de toda duda es que la literatura en el campo se beneficiaría de manera significativa del uso de una terminología consistente y única para referirse al modelo bi-factorial, al modelo de segundo orden y a otros modelos que incluyan uno o varios factores generales.

Por último, es necesario destacar que otros autores han argumentado que la relación subordinada entre los factores de grupo y el factor general debería hacerse explícita en este tipo de modelos. Para ello, el modelo bi-factorial S-1 ha sido propuesto (Eid et al., 2016, 2018). El modelo S-1 incluye un grupo de ítems, correspondientes al mismo factor de grupo, que actúan como indicadores puros del factor general. De esta manera, el factor general pasa a representar varianza específica al contenido del factor de grupo que desaparece (que pasa a ser una faceta de referencia). El modelo S-1 representa una nueva e interesante perspectiva que debería ser explorada en detalle en futuras aplicaciones.

C.2.2 Preguntas sin Resolver en el Modelo Bi-factorial Exploratorio

Tal y como Reise et al. (2018) reflejaron, existen en la actualidad varios mitos relativos al modelo bi-factorial. Entre ellos, su naturaleza exclusivamente confirmatoria, la necesidad de que represente una estructura simple y, por último, el hecho de que los factores de grupos tengan que ser obligatoriamente ortogonales entre sí. En esta tesis se han discutido las dos primeras cuestiones con detalle. Respecto a la tercera, es importante destacar que en análisis factorial, las soluciones oblicuas deben ser preferidas a las ortogonales, ya que permiten la recuperación de la estructura simple en caso de que los factores estén correlacionados, además de proporcionar información adicional valiosa al investigador (Browne, 2001). En este sentido, Jennrich and Bentler (2012) desarrollaron bi-geomin oblicuo para estimar estructuras bi-factoriales con correlaciones entre los factores de grupos. Estos modelos pueden observarse, por ejemplo, cuando existan factores de método (como resultado de los efectos de la aquiescencia o efectos de contenido) o cuando los modelos de bi-factor se apliquen en el contexto del análisis multirasgo-multimétodo o interjueces (Eid et al., 2018), entre otros casos. Además, el uso del modelo de factor S-1 oblicuo podría evitar problemas de identificación (Lorenzo-Seva and Ferrando, 2019b). Así pues, los modelos bi-factoriales oblicuos podrían ser de interés en determinados contextos.

No obstante, los investigadores deben ser cautelosos al aplicar este tipo de modelos oblicuos. En primer lugar, ni siquiera en el caso del análisis factorial general se ha dado una respuesta clara a la pregunta de qué rotación debe utilizarse para estimar correctamente la matriz de correlaciones entre factores. Mientras que algunos autores son partidarios de la rotación geomin (Celimli Alkoy, 2017), recientes avances en la rotación target para modelos oblicuos (Zhang et al., 2018a) no han sido explorados en profundidad como para ser descartados. Además, señalar que aún no se ha presentado una comparación de estos métodos en el contexto de los modelos de factores bi-factoriales. En segundo lugar, asegurar una estimación de calidad de las correlaciones factoriales en el caso bi-factorial no es una tarea sencilla (Lorenzo-Seva and Ferrando, 2019b), ya que frecuentemente da lugar a dificultades en cuanto a su convergencia y a la posibilidad de replicar dichos modelos. En tercer lugar, y de manera más relevante, la introducción de factores de grupo oblicuos hace imposible que los investigadores puedan interpretar los modelos de bi-factor de la misma manera que en el caso ortogonal. En el modelo bi-factorial ortogonal, los factores de grupo se especifican para reflejar variación residual única al factor general, lo que ya no es cierto si los factores de grupo correlacionan entre sí. En palabras de Reise et al. (2018) "[respecto a los modelos bi-factoriales oblicuos] Esto no implica de ninguna manera que la estimación de los modelos de bi-factor con factores de grupo correlacionados no sea complicada o que las soluciones sean fácilmente interpretables; no lo son " (pág. 684, nota al pie 3). Tal escepticismo podría ser la razón por la cual los modelos bi-factoriales oblicuos han recibido tan poca atención hasta ahora en la literatura bi-factorial.

Por último, hay que destacar que todos los modelos bi-factoriales revisados presentan una restricción adicional necesaria, presente incluso en los modelos bi-factoriales exploratorios oblicuos: la ortogonalidad entre los factores generales y de grupo (Eid et al., 2016; Markon, 2019). Esta restricción es necesaria para que el modelo esté identificado y sea estimable (Markon, 2019): "Es probable que las correlaciones entre los factores específicos [i.e., factores de grupo] y el factor general sean inadmisibles independientemente del escenario" (p.12.10). En consecuencia, los modelos bi-factoriales sin restricciones, donde el factor general presenta efectos directos sobre los ítems y los factores de grupo (Yung et al., 1999), podrían considerarse más como una curiosidad matemática que un modelo adecuado para ser aplicado en el mundo real. Sin embargo, como las limitaciones de hoy podrían representar las oportunidades de mañana, el uso y la aproximación a estos modelos debería ser explorada en profundidad en investigaciones futuras.

C.2.3 El Modelo Bi-factorial Expandido

El modelo bi-factorial clásico se ha visto ampliado recientemente para dar cabida a la presencia de fuentes alternativas de variación más allá de un factor general y un factor de grupo sustantivo: (a) los modelos de factores de dos niveles, que presentan un factor general adicional (Cai, 2010; Cai et al., 2011); y (b) el modelo de tri-factorial, donde cada ítem pesa en varios factores de grupo sustantivos simultáneamente (Bauer et al., 2013; Jeon et al., 2018). Así, todo apunta a que la amalgama de modelos bi-factoriales alternativos que podrían derivarse en un futuro cercano únicamente podría crecer exponencialmente. El desarrollo de nuevos modelos alternativos constituye un campo de investigación activo que está ganando adeptos rápidamente.

Sin embargo, los investigadores deben proceder con extrema cautela al aplicar modelos bi-factoriales complejos, comprendiendo sus limitaciones, inconvenientes y cuál es la interpretación específica y teórica de cada modelo. Curiosamente, aún está por verse cómo integrar estos modelos dentro de las taxonomías de modelos bi-factoriales y jerárquicos previamente mencionadas. Por último, es importante tener presente que cuanto mayor sea la complejidad de un modelo, mayor será el número de restricciones necesarias para identificarlo, y mayores serán las posibilidades de que esas restricciones sean incorrectas en una muestra dada. Debido a su naturaleza altamente estructurada, estos modelos se basan en enfoques confirmatorios en los que a menudo se exige al investigador que imponga restricciones clave para asegurar su estimación (i.e., en un modelo de dos niveles, los factores generales pueden correlacionarse entre sí, pero no con los indicadores de factores de grupo; Cai, 2010). Así pues, dichos modelos alternativos podrían beneficiarse de ser considerados desde una perspectiva exploratoria, de tal modo que los investigadores puedan entender si estas herramientas son realmente adecuados a sus datos o no.

C.2.4 La Plausibilidad del Modelo Bi-factorial

Uno de los principales beneficios del modelo bi-factorial es que permite a los investigadores comprender la verosimilitud de la presencia de un factor general subyacente a los datos. Como tal, esta tesis doctoral (Capítulo 5) se centró en el estudio del estadístico omega jerárquico. Este índice desempeña actualmente un papel crucial para comprender la calidad de las puntuaciones totales derivadas de un modelo bi-factorial. Así, omega jerárquico es rutinariamente reportado en la mayoría de las publicaciones que incluyen un modelo bi-factorial. Por ejemplo, este estadístico ha sido un punto focal del debate sobre si un factor general de la personalidad es razonable o no (Arias et al., 2018; Revelle and Wilt, 2013). Por lo tanto, es esperable que este tipo de estadísticos secundarios (incluyendo otros como la varianza común explicada o el índice H) solo crezcan en importancia en el futuro a medida que los modelos bi-factoriales sean aplicados en una mayor cantidad de contextos de investigación (Chen and Zhang, 2018; Reise et al., 2018; Rodriguez et al., 2016).

Lamentablemente, la evaluación de los factores de grupo está muy por detrás de la del factor general, tanto en términos de sofisticación como de calidad de los índices disponibles. Así, es imperativo desarrollar una alternativa apropiada a dichos índices para los factores de grupo en la investigación en los modelos bi-factoriales exploratorios, particularmente dados los desafíos observados en la estimación, interpretación y evaluación de los pesos factoriales y puntuaciones asociadas a estos factores (Bonifay et al., 2017; Markon, 2019; Reise et al., 2013). En la actualidad, existen preocupaciones legítimas tanto sobre el papel que los factores de grupo juegan dentro de los modelos bi-factores como sobre el significado de la varianza que explican (Bonifay et al., 2017; Sellbom and Tellegen, 2019). Por ejemplo, en un modelo bi-factor ortogonal, por definición, un factor de grupo refleja una fuente de varianza residual única y distinta del factor general. Sin embargo, en muchas ocasiones existen serias dudas teóricas y empíricas de la veracidad de este supuesto. Desafortunadamente, la evaluación de la relación entre los factores generales y los de grupo es todavía una tarea desafiante (Eid et al., 2018; Markon, 2019). Sin embargo, sin avances significativos en dicho aspecto, la validez discriminante y convergente del modelo bi-factorial podría verse comprometida.

En cualquier caso, obtener estimaciones fiables de los parámetros de los factores de grupo no es una tarea sencilla. Los factores de grupo suelen definirse con un número pequeño de ítems que presentan una amalgama de pesos factoriales de diferente magnitud. Esta situación hace que las puntuaciones de los factores de grupo sean inestables y poco fiables. En su revisión sistemática, Rodriguez et al. (2015) encontraron que mientras que el factor general estaba, de media, representado con casi 20 ítems, los factores de grupo se estimaron usando un promedio de 7 ítems. Además, mientras que el promedio del omega jerárquico era de 0,80, el promedio del omega jerárquico subescala (i.e., omega jerárquico calculado para el factor de grupo) era de únicamente 0,27. Evidencia similar se ha encontrado en otras revisiones centradas en las escalas de psicopatología (Constantinou and Fonagy, 2019). Así, la cuestión de la utilidad de los factores de grupo más allá de las puntuaciones factoriales generales se ha debatido en la literatura (Bonifay et al., 2017; Rodriguez et al., 2016).

Además, la estimación de los factores de grupo podría verse dificultada por una consecuencia más sutil de la presencia del factor general: la tendencia de dicho factor (dentro de un modelo bi-factorial) a acomodar patrones de respuestas sin sentido (Reise et al., 2016; Watt et al., 2019). Así, se tiene evidencia de que un factor general tiende a absorber una varianza común tanto sustancial como no sustancial independientemente de la intención del investigador. Por ejemplo, la presencia de un fuerte factor del método subyacente a una mayoría de ítems de una escala (p.or ejemplo, un factor de aquiescencia) podría alterar el significado del factor general encontrado. Si bien esta situación abre la posibilidad de que los modelos bi-factoriales sean útiles para controlar los factores de esos métodos, también plantea interrogantes sobre lo que representa el factor general cuando los mismos se pueden controlar explícitamente. De manera similar, se ha argumentado que el modelo bi-factorial tiende a presentar un mejor ajuste que modelos alternativos en ocasiones donde no debería ser así (Bonifay and Cai, 2017; Canivez, 2016; Cucina and Byle, 2017; Murray and Johnson, 2013; Rodriguez et al., 2016). Las razones de dicho efecto podrían estar relacionadas con las restricciones de rango presentes en modelos alternativos como el modelo de segundo orden (Gignac, 2016; Molenaar, 2016; Yang et al., 2017) y la propia (cuestionable) capacidad del modelo bi-factorial para acomodar patrones de datos inverosímiles. Dado que dichas cuestiones siguen sin resolverse, constituyen atractivas líneas de trabajo.

En este punto, se debe realizar una recomendación de precaución respecto al uso generalizado de modelos que incluyen un factor general en contextos donde su estatus teórico es, por decir cuanto menos, dudoso (Borsboom et al., 2003; van Bork et al., 2017; van der Maas et al., 2006). Los investigadores deben entender que, bajo las circunstancias comunes encontradas en evaluación psicológica, la hipótesis de la "variedad positiva"[2] tiende a observarse. La hipótesis de la variedad positiva se refiere al hecho de que los test e ítems que reflejan variables psicológicas tienden a correlacionar de manera positiva independientemente de las características de la aplicación. Este sorprendente hecho fue discutido desde los primeros días del análisis factorial, incluso por el mismo Spearman (Spearman, 1904; van der Maas et al., 2006).

La consecuencia de la presencia de la variedad positiva es que un investigador podría observar un factor general con pesos sustantivos y positivos independientemente del verdadero mecanismo generador de dichas correlaciones. En efecto, extraer un factor general de un conjunto de correlaciones positivas es únicamente una prueba necesaria, pero no suficiente, de su existencia, como lo demuestra el conocido teorema de Perron-Frobenius: para una matriz de varianzas-covarianzas cuyas entradas sean positivas siempre existe un único autovalor real y positivo que corresponde a un autovector de valores positivos. Además, el valor absoluto de este autovalor es el mayor entre todos los autovalores de la matriz [3]. En otras palabras, bajo una matriz de varianzas-covarianzas positivas, siempre se va a encontrar un factor con pesos factoriales positivos, independientemente de qué genere dichas correlaciones positivas. A pesar de que el teorema de Perron-Frobenius se ha utilizado para justificar falta de falsabilidad del modelo factorial general (van Bork et al., 2017), Borg (2018) mostró que la variedad positiva asegura encontrar un factor general de estas características si la dimensionalidad del espacio común es, como mucho, dos. Bajo espacios comunes más grandes, la variedad positiva deja por lo tanto de ser una condición suficiente y pasa a ser una condición necesaria para que se produzca un factor general con las cargas positivas y sustanciales.

Al final, estas consideraciones técnicas deben servir de advertencia para toda decisión en la que se evalúe el uso de factores general se base en fundamentos estadísticos y teóricos sólidos, y no exclusivamente en los resultados de un determinado análisis factorial. En

[2]En castellano, y más precisamente en topología matemática, se denomina variedad a cualquier espacio hipotético geométrico de *n* dimensiones.

[3]para su adaptación al análisis del factor de estructura simple único y jerárquico, véase los Teoremas 1 y 2 en Krijnen (2004).

palabras del propio Thurstone (1947): "A menudo no se comprende la naturaleza exploratoria del análisis factorial. El análisis factorial tiene su principal utilidad en la frontera de la ciencia" (p.56)[4].De hecho, pretender que el análisis factorial vaya más allá de estas fronteras no solo podría ser una tarea poco fructífera y excesívamente optimista [5].

C.2.5 Nuevos Retos en Rotación Factorial

Los resultados de esta tesis doctoral han apoyado firmemente la consideración de la rotación target iterativa parcialmente especificada como un método fiable para aproximar estructuras como el modelo bi-factorial. Como tal, un resultado principal de esta disertación ha sido el desarrollo del algoritmo SLiD, que incluía un método novedoso para encontrar puntos de corte empíricos y específicos para definir la matriz target. Se demostró que este nuevo método resultaba en una mejor recuperación factorial en comparación con los esquemas tradicionales de utilización de puntos de corte únicos. Desafortunadamente, aunque existen varios enfoques diferentes para realizar una rotación target empírica parcialmente especificada, tales como Promaj (Trendafilov, 1994) o Promin (Lorenzo-seva, 1999; Lorenzo-Seva and Ferrando, 2019a), por mencionar unos pocos, estos han sido escasamente comparados en la literatura general, y aún de manera más limitada en el contexto del análisis bi-factorial exploratorio. En este punto, es necesario destacar que durante las últimas etapas de esta tesis se publicó un nuevo algoritmo para estimar modelos bi-factoriales llamado PEBI (Lorenzo-Seva and Ferrando, 2019b). El algoritmo emplea una definición de punto de corte basada en Promin para distinguir entre los pesos sustantivos y cercanos a cero en la matriz target. PEBI es un enfoque prometedor alternativo a SLiD, y que amplía las capacidades de los métodos de rotación target parcialmente especificada al caso de factores de grupo oblicuos o de único factor de grupo. Únicamente cabe aquí animar a los lectores a probar PEBI (así como las otras fantásticas opciones) que se encuentran disponible en el fantástico programa FACTOR (Ferrando and Lorenzo-Seva, 2017b). El autor de esta tesis recomienda encarecidamente a los lectores interesados su uso y consideración.

La rotación target es un área de investigación apasionante dentro de la psicometría. Así, cabe esperar que la rotación target constituya la piedra angular de futuros enfoques exploratorios para aproximar estructuras factoriales complejas (como los modelos de dos o

[4]Thurstone demostró varios de sus avances en análisis factorial exploratorio utilizando medidas físicas de cilindros o cajas. Desafortunadamente, dicha tradición se ha perdido. Este cambio ejemplifica como el trasfondo del análisis factorial ha pasado de las matemáticas a campos de aplicación directa. Eso sí, muchas veces sin la necesaria compresión de las propiedades estadísticas del mismo para su aplicación generalizada

[5]Algunos autores relevantes mostraron una actitud menos cauta al hablar de las posbilidades del análisis factorial: "[...] la existencia de distintos grupos de variables en las configuraciones correlacionales de datos empíricos no representa un accidente más allá de la existencia de diferentes nébulas o la Vía Láctea en el cielo. Las leyes de la naturaleza han generado dichas estructuras en ambos casos." (Cattell, 1978, p.105)

tres factores). Aun así, los nuevos desarrollos que permiten ampliar los métodos de rotación target a nuevas estructuras han tenido un éxito limitado. Por lo tanto, se encuentran a la espera de que los investigadores interesados los estudien en detalle (dentro o fuera del contexto de los modelos bi-factorial): desde las consecuencias de la especificación y uso de targets en otras partes del modelo factorial (Zhang et al., 2018b) hasta la aplicación de técnicas de bootstrapping para comprender la estabilidad de la estimación de la rotación (Paunonen, 1997) así como para desarrollar nuevos esquemas para el refinamiento de la rotación de los objetivos (Lorenzo-seva and Ferrando, 2018). Pero, antes de avanzar en cualquier nueva área de aplicación de la rotación target se debe prestar una atención más detallada a la comparación de los métodos target ya disponibles y ampliamente utilizados: la rotación target completamente y la parcialmente especificada. Dado que la decisión de favorecer una u otra forma podría determinar las futuras líneas de investigación en la materia, las ventajas e inconvenientes deben ser examinadas de manera más cuidadosa en estudios de simulación y los empíricos.

Por último, hay que tener en cuenta que la rotación target no es más que una pequeña porción de un conjunto más amplio de modelos matemáticos que forman parte de la familia del análisis procrusteano generalizado (Crosilla et al., 2019). Así, si bien algunos de los avances más importantes en materia de rotación de objetivos se han realizado en el ámbito de la psicometría, también se han producido contribuciones fundamentales en campos de investigación como el análisis de geométrico de formas, la visión por computadora, etc. (Gower and Dijksterhuis, 2004). Como la conexión entre dichas áreas de investigación ha sido bastante escasa. Como ejemplo, el gran número de alternativas para llevar a cabo la rotación de objetivos parcialmente especificados disponibles en Crosilla (2019; p.20) siguen siendo en gran medida ignorados por nuestra comunidad de investigación.

El interés en las soluciones exploratorias está fomentando los importantes avances que vemos hoy en rotación factorial, donde nuevos enfoques siguen siendo propuestos en las principales revistas y publicaciones del campo. Aunque algunos autores podrían creer que este tema está en gran parte resuelto (Mulaik, 1986), existen todavía muchas áreas de mejora y de investigación. Primero, comprender mejor el funcionamiento del Algoritmo de Proyección de Gradiente. Por ejemplo, confirmando su comportamiento estadístico cuando se compara con otros algoritmos de optimización (Weide and Beauducel, 2019). En este sentido, se sabe poco sobre la sensibilidad de los diferentes parámetros fijados en dicho algoritmo (la tasa de aprendizaje, etc.). Por ejemplo, el uso de matrices de transformación oblicua aleatoria (en lugar de ortogonales, como sugiere Mulaik, 2010, pág. 363) se ha confirmado que tiene un impacto sustancial en el proceso de minimización en análisis no reportados en estos capítulos pero realizados por el autor de esta tesis doctoral. Segundo, explorar algoritmos alternativos de minimización que permitan mejorar la estimación de la rotación en criterios complejos (como el bi-geomin). O para comparar el uso de soluciones algebraicas en la rotación target frente a la solución encontrada por el algoritmo GPA. Tercero, desarrollar rotaciones no centradas en la búsqueda de estructuras simples o la minimización directa o indirecta del número de hiperplanos. Por ejemplo, la propuesta de Jennrich (2004b, 2006) del Criterio de Pérdida de Componentes, que tiene por objeto reducir al mínimo el valor absoluto de un peso factorial determinado sin una referencia directa a los pesos de su fila o su columna, constituye una alternativa interesante que debería explorarse en futuras investigaciones sobre los modelos bi-factoriales. Por el contrario, entender las rotaciones como Varimin (Ertel, 2013), que tienen como objetivo maximizar la complejidad, permitiría a los investigadores entender los límites de los procedimientos de rotación. Cuarto, asegurar la identificación de las soluciones rotadas mediante el empleo de estrategias basadas en el uso de la matriz de información de Fisher (Asparouhov and Muthén, 2009) o similar.

Por último, la rotación factorial podría estar contemplando sus últimos días cómo método principal para identificar una solución factorial y asegurar su simplicidad, al menos en su forma actual. De manera similar a lo acaecido con la transición de los métodos de rotación indirecta a los métodos de rotación directa, una plétora de nuevos métodos alternativos para llevar a cabo estos análisis está ganando terreno en la psicometría convencional. En primer lugar, los enfoques bayesianos basados en el uso de distribuciones previas de pequeña amplitud de manera semiconfirmatoria se están considerando con fuerza como una alternativa a los métodos de rotación para aproximar estructuras complejas (Asparouhov and Muthén, 2009). Sin embargo, aún no está claro qué método debería ser preferible en qué condiciones, particularmente dada la sensibilidad de la aproximación bayesiana a los problemas de especificación errónea (Guo et al., 2019; Marsh et al., 2009; Xiao et al., 2019). Además, ambos enfoques podrían beneficiarse el uno del otro, e incluso llegar a considerarse como complementarios (Moore et al., 2015). En segundo lugar, nuevos métodos de aprendizaje automático basados en la regularización de parámetros [6] han comenzado a aparecer en el campo (Scharf and Nestler, 2019b; Yamamoto et al., 2017). Este enfoque, que de nuevo representa otra perspectiva para resolver el problema de la indeterminación rotacional, tiene varios beneficios, como que la penalización sea un parámetro a estimar o evitar los efectos del sobreajuste. De este modo, la regularización podría representar una perspectiva digna de ser explorada en el futuro. Por último, ha habido algunas reformulaciones verdaderamente interesantes del propio modelo factorial per se, destinadas a superar algunas de las limitaciones

[6] Para los lectores interesados, la rotación de la pérdida de componentes de Jennrich está intrínsecamente relacionada con los métodos de regularización propuestos para llevar a cabo análisis factorial exploratorio Scharf and Nestler (2019a).

presentes en la propuesta original, y considerar algunas de las indeterminaciones revisadas en esta tesis (Adachi and Trendafilov, 2018, 2019; Sočan, 2003; Stegeman, 2016).

C.3 Conclusiones

Esta tesis doctoral debe concluirse con el mismo espíritu con el que se inició, que es aportar una visión crítica de uno de los avances más influyentes de la historia de la psicometría: el modelo bi-factorial. A día de hoy, no debería sorprender a nadie que el modelo bi-factorial constituya un modelo imprescindible dentro de la caja de herramientas de los psicómetras y analistas de datos en psicología. Como tal, se espera que el modelo bi-factorial continúe creciendo en importancia, desempeñando un papel innegable en la configuración del futuro de las principales áreas de la ciencia psicológica. Por lo tanto, es necesario asegurarse de sus puntos fuertes y sus limitaciones se comprenden adecuadamente si se quiere hacer un uso pertinente de esta herramienta estadística. En el fondo, la tesis doctoral se dedicó a profundizar en nuestro conocimiento de este modelo. Y al hacerlo, se espera no solo haber realizado contribuciones limitadas pero significativas en dicho campo, sino también haber inspirado a otros a seguir investigando sobre la naturaleza y las contribuciones del modelo bi-factorial.